D.W. Stroock

An Introduction to the Theory of Large Deviations

Universitext

Springer-Verlag
New York Heidelberg Berlin

Universitext

Editors

F.W. Gehring
P.R. Halmos
C.C. Moore

Universitext

Editors: F. W. Gehring, P.R. Halmos, C.C. Moore

Chern: Complex Manifolds Without Potential Theory
Chorin/Marsden: A Mathematical Introduction to Fluid Mechanics
Cohn: A Classical Invitation to Algebraic Numbers and Class Fields
Curtis: Matrix Groups
van Dalen: Logic and Structure
Devlin: Fundamentals of Contemporary Set Theory
Edwards: A Formal Background to Mathematics I a/b
Edwards: A Formal Background to Higher Mathematics II a/b
Endler: Valuation Theory
Frauenthal: Mathematical Modeling in Epidemiology
Gardiner: A First Course in Group Theory
Godbillon: Dynamical Systems on Surfaces
Greub: Multilinear Algebra
Hermes: Introduction to Mathematical Logic
Kalbfleish: Probability and Statistical Inference I/II
Kelly/Matthews: The Non-Euclidian, The Hyperbolic Plane
Kostrikin: Introduction to Algebra
Luecking/Rubel: Complex Analysis: A Functional Analysis Approach
Lu: Singularity Theory and an Introduction to Catastrophe Theory
Marcus: Number Fields
Meyer: Essential Mathematics for Applied Fields
Moise: Introductory Problem Course in Analysis and Topology
Rees: Notes on Geometry
Reisel: Elementary Theory of Metric Spaces
Rey: Introduction to Robust and Quasi-Robust Statistical Methods (in prep)
Rickart: Natural Function Algebras
Schreiber: Differential Forms
Stroock: An Introduction to the Theory of Large Deviations
Tolle: Optimization Methods

D. W. Stroock

An Introduction to the Theory of Large Deviations

Springer-Verlag
New York Berlin Heidelberg Tokyo

D.W. Stroock
Department of Mathematics
Massachusetts Institute of
Technology
Cambridge, MA 02139
U.S.A.

AMS Classification: 60F10

Library of Congress Cataloging in Publication Data
Stroock, Daniel W.
An introduction to the theory of large deviations.(Universitext)
Bibliography: p.
1. Large deviations. I. Title.
QA273.67.S77 1984 519.5'34 84-10611

9 8 7 6 5 4 3 2 1

ISBN-13: 978-0-387-96021-0 e-ISBN-13: 978-1-4613-8514-1
DOI: 10.1007/978-1-4613-8514-1

Preface

These notes are based on a course which I gave during the academic year 1983-84 at the University of Colorado. My intention was to provide both my audience as well as myself with an introduction to the theory of large deviations .

The organization of sections 1) through 3) owes something to chance and a great deal to the excellent set of notes written by R. Azencott for the course which he gave in 1978 at Saint-Flour (cf. Springer Lecture Notes in Mathematics 774). To be more precise: it is chance that I was around N.Y.U. at the time when M. Schilder wrote his thesis, and so it may be considered chance that I chose to use his result as a jumping off point; with only minor variations, everything else in these sections is taken from Azencott. In particular, section 3) is little more than a rewrite of his exoposition of the Cramer theory via the ideas of Bahadur and Zabel. Furthermore, the brief treatment which I have given to the Ventsel-Freidlin theory in section 4) is again based on Azencott's ideas. All in all, the biggest difference between his and my exposition of these topics is the language in which we have written. However, another major difference must be mentioned: his bibliography is extensive and constitutes a fine introduction to the available literature, mine shares neither of these attributes.

Starting with section 5), I attempted to explain some of the relatively recent advances made by M. Donsker and S.R.S. Varadhan in the theory of large deviations from ergodic phenomena (cf. [D.&V., Parts I & III]). My goal was to see if I could present their theory along the lines suggested by M. Kac in the heuristic discussion given by him in [Kac]. What I found is that the approach proposed by Kac is very closely related to the one successfully employed by Bahadur and Zabel in their work on Sanov-type theorems and that, after some appropriate modifications, their techniques could be made to go quite

far. My efforts in this direction are the contents of sections 5) to 7). In section 8), I abandoned the approach taken in 5) to 7) and returned to the ideas underlying the original paper by Donsker and Varadhan [D.&V.,Oxford] about this subject. Although this approach is restricted to time-reversible processes, I felt that it is the one best suited for possible applications to infinite dimensional situations. Finally, in the course of my studies, I became increasingly aware that there is an interesting relationship between this theory and that of logarithmic Sobolev inequalities. Section 8) is devoted to a somewhat random presentation of my ideas on this relationship.

It is a pleasure to thank the people who helped me prepare these notes. A long-distance but essential role was played by "the grand-master of large deviations", my friend S.R.S. Varadhan. He not only discussed the material with me on several occasions but also sent me copies of the notes he was preparing for his C.B.M.S. lectures. (His C.B.M.S. notes have appeared and cover a great deal of material not treated anywhere else outside of journal articles.) A less appealing but equally essential role was played by H. Heiss and L. Clemens who not only suffered through the delivery of my lectures but also had the stamina to read the typed version of them.

Contents

0. Introduction:

Consider the following situation. E is a Polish space and $\{\mu_\varepsilon : \varepsilon > 0\}$ is a family of probability measures on E such that $\mu_\varepsilon \Rightarrow \delta_{x_o}$ as $\varepsilon \downarrow 0$ (i.e., μ_ε converges weakly to the unit mass at x_o). The study of large deviations is the study of how fast $\mu_\varepsilon(\Gamma) \to 0$ for $\Gamma \in \mathcal{B}_E$ such that $x_o \notin \overline{\Gamma}$. In particular, we will be studying situations in which this convergence is exponentially fast and we will be seeking expressions for

$$- \lim_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma) .$$

We begin with some heuristic observations. First, suppose that all of the μ_ε's are absolutely continuous with respect to some reference measure λ. Then, the facts that the μ_ε's are all probability measures and that they are becoming more and more concentrated at x_o lead one to write $\mu_\varepsilon(dy) = c_\varepsilon \exp(-\frac{1}{\varepsilon} I(y))\, \lambda\, (dy)$, where $I : E \to [0,\infty) \cup \{\infty\}$ and $I(y) = 0$ if and only if $y = x_o$. Assuming that $\varepsilon \log c_\varepsilon \to 0$ as $\varepsilon \downarrow 0$, we then have:

$$\begin{aligned}\lim_{\varepsilon \downarrow 0} \varepsilon \log \mu_\varepsilon(\Gamma) &= \lim_{\varepsilon \downarrow 0} \log \Big(\int_\Gamma \exp(-\tfrac{1}{\varepsilon} I)d\lambda\Big)^{\varepsilon} \\ &= \log \big(\operatorname*{ess\,sup}_{y\in\Gamma}(e^{-I(y)})\big) = -\operatorname*{ess\,inf}_{y\in\Gamma} I(y) ;\end{aligned}$$

here the "ess" refers to λ and we have used the well known fact that if ν is a finite measure, then $\|f\|_{L^p(\nu)} \to \|f\|_{L^\infty(\nu)}$ as $p \to \infty$. Thus, for example, if $E = R^1$ and $\mu_\varepsilon(dy) = (2\pi/\varepsilon)^{-1/2} \exp(-y^2/2\varepsilon)dy$, then

$$(0.1)\quad \lim_{\varepsilon\downarrow 0} \varepsilon \log \mu_\varepsilon (\Gamma) = -\operatorname{ess\,inf}_{y\in\Gamma} y^2/2.$$

Although the preceeding indicates the general structure of the asymptotics in which we are interested, it fails to take into account what to expect when there is no reference measure λ. For example, if no such λ exists, what does "ess inf" mean? To understand what to do in such situations, let us see what we can say about our example without any references to Lebesgue measure. To this end, first suppose that Γ is an open set G. Then $\inf_{y\in G} y^2/2 = \operatorname{ess\,inf}_{y\in G} y^2/2$, and so (0.1) continues to hold for open sets after "ess inf" is replaced by "inf . On the other hand, if Γ is a closed set F, then

$$\lim_{\varepsilon\downarrow 0} \varepsilon \log \mu_\varepsilon(F) = -\operatorname{ess\,inf}_{y\in F} y^2/2 \le -\inf_{y\in F} y^2/2.$$

More generally, what we will be seeking is a statement of the form·

$$\varliminf_{\varepsilon\downarrow 0} \varepsilon \log \mu_\varepsilon(G) \ge -\inf_{y\in G} I(y)$$

for all open sets G, and

$$\varlimsup_{\varepsilon\downarrow 0} \varepsilon \log \mu_\varepsilon(F) \le -\inf_{y\in F} I(y)$$

for all closed sets F. Such a statement is called a <u>large deviation principle</u>.

1. Brownian Motion in Small Time, Strassen's Iterated Logarithm

We are now going to repeat the computation earried out in the preceding, only this time we will be dealing with Wiener measures on path space instead

of Gauss measure on the line. We need some preliminaries.

Let $\Omega = C([0,\infty);R^d)$ and endow Ω with the topology of uniform convergence on compacts. (Clearly, this makes Ω into a Polish space.) Let $\mathcal{M}$ denote the Borel field over Ω. For each $t \geq 0$ and $\omega \in \Omega$, $x(t,\omega)$ denotes the position (value) of ω at time t. Set $\mathcal{M}_t = \sigma(x(s) : 0 \leq s \leq t)$ (i.e., the smallest σ-algebra over Ω with respect to which each of the maps $\omega \to x(s,\omega)$, $0 \leq s \leq t$, is measurable). Clearly $\mathcal{M}_s \subseteq \mathcal{M}_t \subseteq \mathcal{M}$ for all $0 \leq s \leq t$ and $\mathcal{M} = \sigma(\bigcup_{t\geq 0} \mathcal{M}_t)$.

(1.1) Theorem (Wiener): There is a unique probability measure $\mathcal{W}$ on $(\Omega,\mathcal{M})$ such that

$$(1.2) \qquad E^{\mathcal{W}}[\exp(i \sum_{j=1}^{n} (\theta_j, x(t_j)))]$$

$$= \exp[-1/2 \sum_{j,j'=1}^{n} (\theta_j,\theta_{j'})t_j \wedge t_{j'}]$$

for all $n \geq 1$, $0 < t_1 < \ldots < t_n$, and $\theta_1,\ldots,\theta_n \in R^d$. Moreover, if P is a probability measure on $(\Omega,\mathcal{M})$, then following are equivalent:

i) $P = \mathcal{W}$

ii) $P(x(0) = 0) = 1$ and for all $0 \leq s < t$ and $\Gamma \in \mathcal{B}_{R^d}$;
$P(x(t) \in \Gamma | \mathcal{M}_s) = \int_\Gamma (2\pi(t-s))^{-d/2} \exp(-|y-x(s)|^2/2(t-s))dy$

iii) $P(x(0) = 0) = 1$ and for all $n \geq 1$ and $0 = t_0 < t_1 < \ldots < t_n$, $\{x(t_j) - x(t_{j-1}) : 1 \leq j \leq n\}$ is a family of independent R^d-valued Gaussian random variables, the j^{th} one of which having mean 0 and covariance $(t_j - t_{j-1})I$.

In particular, if $n \geq 1$ and $0 = t_o < \ldots < t_n$, then the σ-algebras $F_j = \sigma(x(t) - x(t_{j-1}) : t_{j-1} \leq t \leq t_j)$ are independent under $\mathcal{W}$.

Proof: The only non-trivial assertion is that $\mathbb{W}$ exists. For a proof, see any text having "Brownian motion" or "Stochastic processes" in its title. ■

It is often convenient to have the following notion. Let $(E,\mathcal{F},P)$ be a probability space and $\{\mathcal{F}_t : t \geq 0\}$ a non-decreasing family of sub σ-algebras of $\mathcal{F}$. Given a map $\beta : [0,\infty) \times E \to R^d$, we say that $(\beta(t),\mathcal{F}_t,P)$ is an (R^d-valued) Brownian motion if

a) for P almost all $q \in E$, $t \to \beta(t,q)$ is continuous; and for all $t \geq 0$, $\beta(t)$ (i.e. $q \to \beta(t,q)$) is $\mathcal{F}_t$ measurable;

b) for all $0 \leq s < t$ and $\Gamma \in \mathcal{B}_{R^d}$,

$$P(\beta(t) \in \Gamma | \mathcal{F}_s) = \int_\Gamma (2\pi(t-s))^{-d/2} \exp(-|y - \beta(s)|^2/2(t-s)) .$$

Clearly, if $\tilde{\mathcal{F}}_t = \sigma(\beta(s) : 0 \leq s \leq t)$, then $(\beta(t),\tilde{\mathcal{F}}_t,P)$ is a Brownian motion if and only if $\beta(\cdot)$ is P-a.s. continuous and $P \circ (\beta(\cdot) - \beta(0))^{-1}$ (i.e. the measure induced on Ω by the map $q \in E' \to (\beta(\cdot,q) - \beta(0,q)) \in \Omega$, where $E' = \{q : t \to \beta(t,q) \text{ is continuous}\}$) is $\mathbb{W}$. Given a probability space $(E,\mathcal{F},P)$ and a $\mathcal{B}_{[0,\infty)} \times \mathcal{F}$-measurable $\beta : [0,\infty) \times E \to R^d$, we will say that $\beta(\cdot)$ is a (R^d-valued) P-Brownian motion if $(\beta(t), \tilde{\mathcal{F}}_t,P)$ is a Brownian motion.

(1.3) Exercises:

1) Suppose that $(\beta(t),\mathcal{F}_t,P)$ is a Brownian motion and that $P(\beta(0) = 0) = 1$. Given $\lambda > 0$, set $\beta_\lambda(t) = \lambda^{-1/2} \beta(\lambda t)$, and $\mathcal{F}_t^\lambda = \mathcal{F}_{\lambda t}$, $t \geq 0$. Show that $(\beta_\lambda(t),\mathcal{F}_t^\lambda,P)$ is a Brownian motion.

2) Let $\beta(\cdot)$ be a P-Brownian motion with $P(\beta(0) = 0) = 1$. Show that $t\beta(1/t) \to 0$, P-a.s., as $t \downarrow 0$. Set $\hat{\beta}(t) = t(\beta(1/t))$. Show that $\hat{\beta}(\cdot)$ is a P-Brownian motion.

We now want to prove the following theorem due to M. Schilder. For

$\varepsilon > 0$, define $X_\varepsilon(t,\omega) = \varepsilon^{1/2}\, x(t,\omega)$, $(t,\omega) \in [0,\infty) \times \Omega$. Set $\mathcal{W}_\varepsilon = \mathcal{W} \circ X_\varepsilon(\cdot)^{-1}$. Clearly $\mathcal{W}_\varepsilon \Rightarrow \delta_0$ as $\varepsilon \downarrow 0$, where δ_0 denotes the unit mass at the path $\phi(t) = 0$, $t \geq 0$. We want to prove a large deviations result for $\{\mathcal{W}_\varepsilon : \varepsilon > 0\}$. More precisely, we will show that if $T > 0$ and F is any closed $\mathcal{M}_T$-measurable subset of Ω , then

$$(1.4) \qquad \overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log \mathcal{W}_\varepsilon(F) \leq -\inf_{\phi\in F} I_T(\phi)$$

where

$$(1.5) \qquad I_T(\phi) = \begin{cases} \infty \text{ if } \phi(0) \neq 0 \text{ or } \phi|_{[0,T]} \text{ is not absolutely continuous} \\ 1/2\int_0^T |\dot{\phi}(t)|^2\, dt \quad \text{if } \phi(0) = 0 \text{ and } \phi|_{[0,T]} \text{ is absolutely continuous.} \end{cases}$$

We also want the complementary inequality:+ if G is an $\mathcal{M}_T$-measurable open set, then

$$(1.6) \qquad \underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log \mathcal{W}_\varepsilon(G) \geq -\inf_{\phi\in G} I_T(\phi) .$$

(1.7) <u>Exercise</u>: Show that the example discussed in section 0 is a special case of (1.4) and (1.6) .

Before turning to a rigorous derivation of Schilder's Theorem, note that the result is intuitively clear. Indeed, pretending that there is a "flat measure" on $C([0,T];\, R^d)$, it is clear that

$$\mathcal{W}_\varepsilon(d\omega) = C_\varepsilon \exp(-1/2\varepsilon \int_0^T |\dot{x}(t,\omega)|^2 dt)\, d\omega \quad \text{on } \mathcal{M}_T .$$

Here C_ε , $\dot{x}(t,\omega)$, and $d\omega$ are all meaningless. Hence, by the intuition used in section (0), (1.4) and (1.6) are just what we should expect.

<u>(1.8)</u> <u>Lemma</u>: As a function on $C([0,\infty);R^d]$ into $[0,\infty) \cup \{\infty\}$, I_T is lower semi-continuous with respect to the semi-norm $\|\cdot\|_T^0$ given by: $\|\phi\|_T^0 = \sup_{0\leq t\leq T} |\phi(t)|$. Moreover, for any $C < \infty$, $\{\phi \in C([0,\infty);R^d:\ I_T(\phi) \leq C\}$ is compact with respect to $\|\cdot\|_T^0$,

<u>Proof</u>: The second assertion is an easy corollary of the Ascoli-Arzela theorem plus the lower semi-continuity of I_T with respect to $\|\cdot\|_T^0$. (Note that $|\psi(t_2) - \psi(t_1)| \le (2I_T(\psi))^{1/2}\,(t_2 - t_1)^{1/2}$ for $0 \le t_1 < t_2 \le T$.)

To prove the lower semi-continuity of I_T , we assume for convenience that $T = 1$. We claim that

$$(1.9) \qquad 2\,I_1(\psi) = \sup_n n \sum_1^n \left| \psi\left(\frac{k}{n}\right) - \psi\left(\frac{k-1}{n}\right)\right|^2 \text{ if } \psi(0) = 0 .$$

Since $n\sum_1^n \left|\psi\left(\frac{k}{n}\right) - \psi\left(\frac{k-1}{n}\right)\right|^2$ is $\|\cdot\|_1^0$-continuous for each n and $\{\psi \in C([0,\infty),R^d) : \psi(0) = 0\}$ is closed, surely (1.9) proves that I_T is $\|\cdot\|_1^0$-lower semi-continuous.

To see that (1.9) holds, set $H_1([0,1];R^d) = \{\psi \in C([0,\infty);R^d): I_1(\psi) < \infty\}$, $J_n(\psi) = n\sum_1^n \left|\psi\left(\frac{k}{n}\right) - \psi\left(\frac{k-1}{n}\right)\right|^2$, and $J(\psi) = \sup_n J_n(\psi)$. We first note that $\psi(0) = 0$ and $J(\psi) < \infty$ imply $\psi \in H_1([0,1];R^d)$. Indeed, let $\psi^{(n)}$ denote the polygonal interpolant of ψ between points $\frac{k}{n}$, $k \ge 0$. Then $J_n(\psi) = \int_0^1 |\dot\psi^{(n)}(t)|^2\,dt$. Hence $J(\psi) < \infty$ implies that $\sup_n \int_0^1 |\dot\psi^{(n)}(t)|^2\,dt = M < \infty$. Thus, for any $\varphi \in C_0^\infty([0,1];R^d)$:
$$\left|\int_0^1 (\psi(t),\dot\varphi(t))dt\right| = \lim_{n\to\infty} \left|\int_0^1 (\psi^{(n)}(t),\dot\varphi(t))dt\right| \le M^{1/2}\,\|\varphi\|_{L^2([0,1];R^d)} .$$
That is, $\psi|_{[0,1]}$ has one distributional derivative $\dot\psi|_{[0,1]} \in L^2([0,1];R^d)$. Thus, $\psi \in H_1([0,1];R^d)$. We therefore only have to prove (1.9) when $\psi \in H_1([0,1];R^d)$. To this end, note that, by Schwartz's inequality $J_n(\psi) \le 2I_1(\psi)$ for all $n \ge 1$ and $\psi \in H_1([0,1];R^d)$. Also, by the triangle inequality, $|J_n(\psi)^{1/2} - J_n(\psi')^{1/2}| \le J_n(\psi - \psi')^{1/2}$ for any $n \ge 1$ and $\psi,\psi' \in C([0,\infty);R^d)$. Thus $|J(\psi)^{1/2} - J(\psi')^{1/2}| \le J(\psi - \psi')^{1/2}$, $\psi,\psi' \in H_1([0,1];R^d)$. In particular, if $\{\psi_n\}_1^\infty \cup \{\psi\} \subseteq H_1([0,1];R^d)$

satisfy $I_1(\phi_n - \phi) \to 0$ as $n \to \infty$, then $J(\phi_n) \to J(\phi)$. It follows that we need prove (1.9) only for a set of ϕ's which are $I_1(\cdot)^{1/2}$-dense in $H_1([0,1];R^d)$. In particular, $\phi \in C^\infty([0,\infty);R^d)$ with $\phi(0) = 0$ will do. But if $\phi \in C^\infty([0,\infty);R^d)$ with $\phi(0) = 0$, then

$$2I_1(\phi) \ge J(\phi) \ge \lim_{n\to\infty} n \sum_1^n \left|\phi\left(\frac{k}{n}\right) - \phi\left(\frac{k-1}{n}\right)\right|^2 = 2\, I_1(\phi).$$

∎

As a consequence of (1.8), we have the following. Given a closed, non-empty, $\mathcal{M}_T$-measurable set F, we have

(1.10) $$\inf_{\phi\in F^\delta} I_T(\phi) \uparrow \inf_{\phi\in F} I_T(\phi)$$

as $\delta \downarrow 0$, where $F^\delta = \{\phi : (\exists\phi' \in F)\ \|\phi - \phi'\|_T^0 < \delta\}$. To prove (1.10), set $\ell_\delta = \inf_{\phi\in F^\delta} I_T(\phi)$ and $\ell = \inf_{\phi\in F} I_T(\phi)$. Clearly $\ell_\delta \uparrow$ as $\delta\downarrow$. Suppose that $\ell_\delta \le \ell' < \ell$ for all $\delta \ge 0$. Choose $\phi_n \in F^{1/n}$ so that $I_T(\phi_n) \le \ell_{1/n} + 1/n \le \ell' + 1/n$. By the second part of (1.8), there is $\|\cdot\|_T^0$-convergent subsequence $\{\phi_n'\}$ of $\{\phi_n\}$. Choose ϕ so that $\|\phi_n' - \phi\|_T^0 \to 0$. Clearly $\phi \in F$. At the same time, $I_T(\phi) \le \varliminf_{n'\to\infty} I_T(\phi_n') \le \ell' < \ell$. Thus we get a contradiction unless (1.10) holds.

(1.11) Lemma: For any $T > 0$ and $\delta > 0$,

$$\mathcal{W}(\|x(\cdot)\|_T^0 \ge \delta) \le 2d \exp(-\delta^2/2dT).$$

Proof: From ii) in Theorem (1.1), we see that, for any $\theta \in R^d$, $X_\theta(t) = \exp((\theta, x(t)) - \frac{|\theta|^2}{2} t)$ is a $(\Omega, \mathcal{M}_t, \mathcal{W})$-martingale. Now fix $\theta \in S^{d-1}$ and apply Doob's inequality to $X_{\lambda\theta}(\cdot)$ for $\lambda > 0$:

$$\mathcal{W}(\sup_{0\le t\le T} (\theta, x(t)) \ge \delta)$$

$$\le \mathcal{W}(\sup_{0\le t\le T} X_{\lambda\theta}(t) \ge \exp(\lambda\delta - \lambda^2 T/2))$$

$$\le \exp(-\lambda\delta + \lambda^2 T/2) \ .$$

Taking $\lambda = \delta/T$, we find that

$$\mathbb{W}(\sup_{0\le t\le T} (\theta, x(t)) \ge \delta) \le e^{-\delta^2/2T} \ .$$

Since,

$$\mathbb{W}(\sup_{0\le t\le T} |x(t)| \ge \delta) \le 2d \sup_{\theta\in S^{d-1}} \mathbb{W}(\sup_{0\le t\le T} (\theta, x(t)) \ge \delta/d^{1/2}) \ ,$$

our estimate follows. ∎

(1.12) Proof of (1.4) :

Let F be given. Without loss of generality, we assume that $F \subseteq \{\psi : \psi(0) = 0\}$ and that $F \neq \phi$.

Given $\psi \in C([0,\infty);R^d)$ and $n \ge 1$, let $\psi^{(n)}$ denote the polygonal interpolant of ψ between times kT/n, $k \ge 0$, and set

$$V_n(\psi) = 2I_T(\psi^{(n)}) \ .$$

Given $\delta > 0$, set $\ell_\delta = \inf_{\psi\in F^\delta} I_T(\psi)$. Then for all $n \ge 1$ and $\delta > 0$;

$$F \subseteq \{\psi : \psi^{(n)} \in F^\delta\} \cup \{\psi : \|\psi - \psi^{(n)}\|_T^0 \ge \delta\}$$

$$\equiv A_n(\delta) \cup B_n(\delta) \ .$$

We first estimate $\mathbb{W}_\varepsilon(A_n(\delta))$. To this end, note that

$$A_n(\delta) \subseteq \{\psi : V_n(\psi) \ge 2\ell_\delta\}$$

and so

$$\mathbb{W}_\varepsilon(A_n(\delta)) \le \mathbb{W}(V_n(x(\cdot)) \ge 2\ell_\delta/\varepsilon) \ .$$

Next, observe that $(\frac{n}{T})^{1/2} (x(\frac{kT}{n}) - x(\frac{(k-1)T}{n}))$, $1 \le k \le n$, are independent, R^d-valued Gaussians with mean 0 and covariance I. Thus

$$V_n(x(\cdot)) = \frac{n}{T}\sum_1^n \left|x(\frac{kT}{n}) - x(\frac{(k-1)T}{2})\right|^2$$

has a χ^2 distribution with parameter nd. In particular,

$$\mathbb{W}(V_n(x(\cdot)) \geq 2\ell_\delta/\varepsilon)$$

$$= c_n\int_{\ell_\delta/\varepsilon}^{\infty} e^{-u/2}\, u^{nd/2-1}\, du$$

$$= e^{-2\ell_\delta/\varepsilon}\; c_n\int_0^{\infty} e^{-u/2}\,(u + 2\ell_{\delta/\varepsilon})^{nd/2-1}\, du\ ;$$

and so for each $n \geq 1$ and $\delta > 0$ there is a $K_n(\delta) < \infty$ such that

$$\mathbb{W}_\varepsilon(A_n(\delta)) \leq (K_n(\delta)/\varepsilon^{nd/2})\; e^{\ell_\delta/\varepsilon}$$

so long as $0 < \varepsilon \leq 1$.

We next turn to $\mathbb{W}_\varepsilon(B_n(\delta))$. But

$$\mathbb{W}_\varepsilon(B_n(\delta)) = \mathbb{W}(B_n(\delta/\varepsilon^{1/2}))$$

$$\leq \mathbb{W}\Big(\max_{0\leq k\leq n-1}\ \sup_{0\leq t\leq T/n} |x(kT/n + t) - x(kT/n)| \geq \delta/2\varepsilon^{1/2}\Big)$$

$$\leq 2nd\,\exp(-n\delta^2/8dT\varepsilon)\ ,$$

where we have used the fact that $\beta(t) = x(s+t) - x(s)$, $t \geq 0$, is a $\mathbb{W}$-Brownian motion for each $s \geq 0$.

Now fix $\delta > 0$ and choose $n \geq 1$ so that $n\delta^2/8dT > \ell_\delta$. Then, from the preceding three paragraphs we have

$$\mathbb{W}_\varepsilon(F) \leq \mathbb{W}_\varepsilon(A_n(\delta)) + \mathbb{W}_\varepsilon(B_n(\delta))$$

$$\leq e^{-\ell_\delta/\varepsilon}[(K_n(\delta)/\varepsilon^{nd/2}) + 2nd\, e^{-\eta/\varepsilon}],$$

where $\eta = n\delta^2/8dT - \ell_\delta > 0$. In particular:

$$\varepsilon\,\log \mathbb{W}_\varepsilon(F) \leq -\ell_\delta + \varepsilon\,\log[K_n(\delta)/\varepsilon^{nd/2} + 2nd\, e^{-\eta/\varepsilon}]$$

and so

$$\overline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log \mathcal{W}_\varepsilon(F) \le - \inf_{\psi \in F^\delta} I_T(\psi) \ .$$

Using (1.10) we get (1.4) by letting $\delta \downarrow 0$. ■

The proof of (1.6) depends on a result which is a great importance in and of itself. We now present this result (originally due to Cameron and Martin) in a more general form than is necessary for our immediate purposes.

(1.13) Theorem: Given $\theta \in R^d$, define $X_\theta : [0,\infty) \times \Omega \to R^1$ by $X_\theta(t) = \exp((\theta, x(t)) - \frac{|\theta|^2}{2} t)$. Given $T > 0$, $P|_{\mathcal{M}_T} = \mathcal{W}|_{\mathcal{M}_T}$ if and only if $(X_\theta(t \wedge T), \mathcal{M}_t, P)$ is a mean one martingale for each $\theta \in R^d$. Moreover, if $\eta : [o,\infty) \times \Omega \to R^d$ is a progressively measurable function with respect to $\{\mathcal{M}_t : t \ge 0\}$ (i.e. for all $T > 0$, $\eta|_{[0,T]\times\Omega}$ is $\mathcal{B}_{[0,T]} \times \mathcal{M}_T$-measurable) satisfying $\sup_\omega \int_0^T |\eta(t,\omega)|^2 \, dt < \infty$ for each $T > 0$ and if

(1.14) $X_\eta(t) = \exp(\int_0^t (\eta(s) , dx(s)) - 1/2 \int_0^t |\eta(s)|^2 ds)$, $t \ge 0$,

then $(X_\eta(t), \mathcal{M}_t, \mathcal{W})$ is a martingale. (The quantity $\int_0^t (\eta(s), dx(s))$ is the Itô stochastic integral of $\eta(\cdot)$.) Finally, given such an η ; set $\beta^\eta(t) = x(t) + \int_0^t \eta(s)\, ds$ and $P^\eta = \mathcal{W} \circ (\beta^\eta(\cdot))^{-1}$. Then; for each $T > 0$, $P^\eta|_{\mathcal{M}_T} << \mathcal{W}|_{\mathcal{M}_T}$ and $\dfrac{dP^\eta|_{\mathcal{M}_T}}{d\mathcal{W}|_{\mathcal{M}_T}} = X_\eta(T)$.

Proof: The characterization of $\mathcal{W}$ in terms of the X_θ's is an easy exercise based on ii) of Theorem (1.1) .

To prove that X_η in (1.14) is a $\mathcal{W}$-martingale , one starts with the case in which η is uniformly bounded and simple (i.e. satisfies $\eta(t) = \eta([Nt]/N)$, $t \ge 0$, for some $N \ge 1$) . The general case is then

obtained by an easy limit procedure. More details can be found in Theorem 4.2.1 of [S. & V.] .

Finally, to prove the last assertion it again suffices to handle the situation in which η is uniformly bounded and simple; the general case then follows after taking limits . Assuming that η is uniformly bounded and simple, note that $\int_0^t(\eta(s), dx(s))$ and therefore $X_\eta(t)$ are defined everywhere on Ω . (Indeed, $\int_0^t(\eta(s), dx(s)) = \sum_k (\eta(\frac{k}{N}) \cdot x(\frac{k+1}{N} \wedge t)$ $-x(\frac{k}{N} \wedge t))$ if $\eta(t) = \eta([Nt]/N)$, $t \geq 0$.) In particular we can define P^η integrals of $X_\eta(t)$ unambiguously. Given $T > 0$, define $Q = 1/X_\eta(T)\, P^\eta$ (i.e. $Q << P^\eta$ and $\frac{dQ}{dP^\eta} = 1/X_\eta(T))$. If we can show that $Q|_{\mathcal{M}_T} = \mathcal{W}|_{\mathcal{M}_T}$, then we will know that $P^\eta|_{\mathcal{M}_T} = X_\eta(T)\mathcal{W}|_{\mathcal{M}_T}$. To this end, let $\theta \in R^d$ $0 \leq t_1 < t_2 \leq T$ and $A \in \mathcal{M}_{t_1}$ be given. Then

$$E^Q[X_\theta(t_2), A] = E^{P^\eta}[X_\theta(t_2)/X_\eta(T), A]$$

$$= E^{\mathcal{W}}[X_{\theta-\eta}(t_2)\, X_{-\eta}(T)/X_{-\eta}(t_2) , A]$$

$$= E^{\mathcal{W}}[X_{\theta-\eta}(t_2), A] = E^{\mathcal{W}}[X_{\theta-\eta}(t_1), A]$$

$$= E^{\mathcal{W}}[X_{\theta-\eta}(t_1)\, X_{-\eta}(T)/X_{-\eta}(t_1), A]$$

$$= E^{P^\eta}[X_\theta(t_1)/X_\eta(T), A] = E^Q[X_\theta(t_1), A] ,$$

where we have made repeated use of the martingale property of exponentials of the form given in (1.14). Thus, we have shown that $(X_\theta(t \wedge T), \mathcal{M}_t, Q)$ is a martingale for all $\theta \in R^d$; and so by the first part, $Q|_{\mathcal{M}_T} = \mathcal{W}|_{\mathcal{M}_T}$. ∎

(1.15) Proof of (1.6):

Let G be an open $\mathcal{M}_T$-measurable subset of Ω satisfying

$G \cap \{\psi : \psi(0) = 0\} \neq \emptyset$. Set $\ell = \inf_{\psi \in G} I_T(\psi)$. Given $\alpha > 0$, we can find $\psi^0 \in G \cap C^2([0,\infty);R^d)$ such that $\psi^0(0) = 0$ and $I_T(\psi^0) \geq \ell - \alpha$. Choose $\delta^0 > 0$ so that $B_T(\psi^0,\delta^0) = \{\psi : \|\psi - \psi^0\|_T^0 < \delta^0\} \subseteq G$. Then for $0 < \delta \leq \delta^0$:

$$\mathcal{W}_\varepsilon(G) \geq \mathcal{W}_\varepsilon(B(\psi^0,\delta))$$

$$= \mathcal{W}(\{\omega : \|\varepsilon^{1/2} x(\cdot,\omega) - \psi^0(\cdot)\|_T^0 < \delta)$$

$$= \mathcal{W}(\{\omega : \|x(\cdot,\omega) - 1/\varepsilon^{1/2} \psi^0(\cdot)\|_T^0 < \delta/\varepsilon^{1/2})$$

$$= P_\varepsilon(\{\omega : \|x(\cdot,\omega)\|_T^0 < \delta/\varepsilon^{1/2})$$

where $P_\varepsilon = \mathcal{W} \circ (\beta^\varepsilon(\cdot))^{-1}$ with $\beta^\varepsilon(t) = x(t) - 1/\varepsilon^{1/2} \psi^0(t)$, $t \geq 0$. By Theorem (1.13) , $P_\varepsilon|_{\mathcal{M}_T} = X_\varepsilon(T)\mathcal{W}|_{\mathcal{M}_T}$, where

$$X_\varepsilon(T) \equiv \exp(- 1/\varepsilon^{1/2} \int_0^T (\dot\psi^0(s), dx(s)) - 1/\varepsilon\, I_T(\psi^0)) .$$

Note that

$$\int_0^T (\dot\psi^0(s), dx(s)) = (\dot\psi^0(T), x(T)) - \int_0^T (\ddot\psi^0(s), x(s))ds$$

$$\leq (1 + T) \|\psi^0\|_{C^2([0,T);R^d)} \|x(\cdot)\|_T^0 .$$

Thus, with $K = (1 + T) \|\psi^0\|_{C^2([0,T);R^d)}$

$$\mathcal{W}_\varepsilon(G) \geq e^{-I_T(\psi^0)/\varepsilon} E^{\mathcal{W}}[e^{-(K\|x(\cdot)\|_T^0/\varepsilon^{1/2})} , \|x(\cdot)\|_T^0 < \delta/\varepsilon^{1/2}]$$

$$\geq \exp(- I_T(\psi^0)/\varepsilon - K\delta/\varepsilon)\mathcal{W}(\|x(\cdot)\|_T^0 < \delta/\varepsilon^{1/2}) .$$

Since, by continuity of the paths $x(\cdot,\omega)$, $\mathcal{W}(\|x(\cdot)\|_T^0 < \delta/\varepsilon^{1/2}) \to 1$, we now see that:

$$\varliminf_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{W}_\varepsilon(G) \geq -I_T(\psi^0) - K\delta$$

$$\geq -\ell - \alpha - K\delta$$

for all $\alpha > 0$ and $0 < \delta \leq \delta^0$. Clearly this implies (1.6) . ∎

We have now proved Schilder's theorem. To summarize, we state:

(1.16) <u>Theorem</u>: For $\varepsilon > 0$, set $X_\varepsilon(t) = \varepsilon^{1/2} x(t)$, $t \geq 0$ and $\mathbb{W}_\varepsilon = \mathbb{W} \circ X_\varepsilon(\cdot)^{-1}$. Let $T > 0$ be given and define I_T by (1.5) . If F is any closed , $\mathcal{M}_T$ -measurable subset of Ω , then

$$\varlimsup_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{W}_\varepsilon(F) \leq - \inf_{\psi \in F} I_T(\psi) .$$

If G is any open , $\mathcal{M}_T$-measurable subset of Ω , then

$$\varliminf_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{W}_\varepsilon(G) \geq - \inf_{\psi \in G} I_T(\psi) .$$

Having worked so hard to get Theorem (1.16) , it is only fair that we demonstrate that there are nice consequences. Perhaps ths most striking application is the beautiful theorem due to V. Strassen on the law of the iterated logarithm.

(1.17) <u>Theorem</u>: Define $\phi(n) = (2n \log_2 n)^{1/2}$ for $n \geq 2$ ($\log_2 x = \log(\log x)$ for $x > 1$) . Given $n \geq 2$ and $T > 0$, set $\xi_n(t,\omega) = x(nt,\omega)/\phi(n)$, $t \in [0,T]$ and $\omega \in \Omega$. Then for $\mathbb{W}$-almost all ω , the sequence $\{\xi_n(\cdot,\omega)\}_2^\infty$ has the following properties:

i) $\{\xi_n(\cdot,\omega)\}_2^\infty$ is precompact in $C([0,T],R^d)$;

ii) if $\{\xi_{n'}(\cdot,\omega)\}_2^\infty$ is a convergent sub-sequence of $\{\xi_n(\cdot,\omega)\}_2^\infty$ and ψ is its limit, then $2I_T(\psi) \leq 1$;

iii) if $\psi \in C([0,T];R^d)$ with $2I_T(\psi) \leq 1$, then there is a subsequence of $\{\xi_n(\cdot,\omega)\}_2^\infty$ which converges to ψ .

In particular, if $\Phi : C([0,T);R^d) \to R^1$ is a continuous function, then

(1.18) $$W(\overline{\lim_{n\to\infty}}\, \Phi(\xi_n(\cdot)) = \sup_{\psi\in K_T} \Phi(\psi)) = 1 ,$$

where $K_T = \{\psi \in C([0,T];R^d) : \psi(0) = 0 \text{ and } 2I_T(\psi) \le 1\}$.

(1.19) Exercise: Given the rest of Theorem (1.17) , prove (1.18) . Apply (1.18) to prove the classical statement that $W(\overline{\lim_{n\to\infty}}\, x_1(n)/\phi(n) = 1) = 1$.

(1.20) Lemma: Let $K \subset\subset C([0,T];R^d)$, $\Psi \in C([0,\infty);R^d)$, and $\Lambda \subseteq (1,\infty)$ satisfy $1 \in \overline{\Lambda}$. For $n \ge 2$, set $\psi_n(t) = \Psi(nt)/\phi(n)$, $0 \le t \le T$. Given $\lambda > 1$, set $n_m(\lambda) = [\lambda^m]$. Assume that for each $\lambda \in \Lambda$, $\overline{\lim_{m\to\infty}}\, \|\psi_{n_m(\lambda)} - K\|_T^0 = 0$. Then the sequence $\{\psi_n\}_2^\infty$ is precompact. Moreover, every convergent subsequence converges to an element of K .

Proof: Since K is compact, we can find $M < \infty$ and a non-decreasing function $\rho : (0,T] \to (0,\infty)$ satisfying $\lim_{t\downarrow 0} \rho(t) = 0$ such that

$$\sup_{\psi\in K} \|\psi\|_T^0 \le M \quad\text{and}\quad \sup_{\psi\in K} |\psi(t) - \psi(s)| \le \rho(t-s) ,\ 0 \le s < t \le T .$$

Given $\delta > 0$, choose $\lambda \in \Lambda$ so that $\rho((1 - 1/\lambda)T) < \delta$ and $(\lambda-1)(M+\delta) < \delta$. Next, choose an integer $L \ge 2$ so that $(\log_2 \lambda x)/\log_2 x \le \lambda$ for $x \ge L$. Finally choose $m(\lambda,\delta)$ so that $\|\psi_{n_m(\lambda)} - K\|_T^0 < \delta$ for $m \ge m(\lambda,\delta)$ and set $n(\delta) = L \vee m(\lambda,\delta)$. Given $n \ge n(\delta)$, choose m so that $\lambda^m \le n \le \lambda^{m+1}$ and set $N = [\lambda^{m+1}]$. Since $m \ge m(\lambda,\delta)$, we know that $\|\psi_N - K\|_T^0 < \delta$. Thus $\|\psi_n - K\|_T^0 < \delta + \|\psi_n - \psi_N\|_T^0$. Noting that

$$\psi_n(t) = \Psi(nt)/\phi(n) = \Psi(\tfrac{n}{N} Nt)/\phi(n)$$

$$= \psi_N(\tfrac{n}{N} t) \frac{\phi(N)}{\phi(n)} ,$$

we see that

$$\|\psi_n - \psi_N\|_T^0 = \sup_{0\le t\le T} \left|\frac{\phi(N)}{\phi(n)} \psi_N(\tfrac{n}{N} t) - \psi_N(t)\right|$$

$$\le \left|\frac{\phi(N)}{\phi(n)} - 1\right| \, \|\psi_N\|_T^0 + \sup_{0\le t\le T} \left|\psi_N\left(\frac{n}{N} t\right) - \psi_N(t)\right|$$

$$\le \left|\frac{\phi(N)}{\phi(n)} - 1\right| (M + \delta) + 2\delta + \rho\left(\left(1 - \frac{n}{N}\right)T\right) .$$

Since $n \le N \le \lambda n$,

$$1 \le \frac{\phi(N)}{\phi(n)} \le \frac{\phi(\lambda n)}{\phi(n)} = \left(\lambda \frac{\log_2 \lambda n}{\log_2 n}\right)^{1/2} \le \lambda .$$

Hence $\left|\frac{\phi(N)}{\phi(n)} - 1\right| (M + \delta) \le (\lambda - 1)(M + \delta) \le \delta$. At the same time, $\rho\left(\left(1 - \frac{n}{N}\right)T\right) \le \rho((1 - 1/\lambda)T) < \delta$. We have therefore shown that $\|\psi_n - K\|_T^0 < 5\delta$ for $n \ge n(\delta)$.

We now know that $\lim_{n\to\infty} \|\psi_n - K\|_T^0 = 0$. Since K is compact, the desired result is easy from here. ■

(1.21) <u>Proof of Theorem (1.17)</u>:

We first prove i) and ii) . To this end, let $\lambda > 1$ be fixed for the moment. Given $\delta > 0$, we can choose $1 < \gamma < \inf_{\psi \notin K_T^\delta} 2I_T(\psi)$. Noting that $\mathcal{W}(\xi_{n_m(\lambda)}(\cdot) \notin K_T^\delta) = \mathcal{W}_{\varepsilon_m(\lambda)}(x(\cdot) \notin K_T^\delta)$, where $\varepsilon_m(\lambda) = (2 \log_2 n_m(\lambda))^{-1}$, we see from (1.4) that $\mathcal{W}(\xi_{n_m(\lambda)}(\cdot) \notin K_T^\delta) \le \exp(-\gamma \log_2 n_m(\lambda)) \le \frac{1}{\log \lambda} \frac{1}{(m-1)^\gamma}$ for sufficiently large m's . Hence, $\sum_m \mathcal{W}(\xi_{n_m(\lambda)}(\cdot) \notin K_T^\delta) < \infty$; and so $\mathcal{W}((\exists N)(\forall m \ge N)\ \xi_{n_m(\lambda)}(\cdot) \in K_T^\delta) = 1.$ Since this is true for every $\delta > 0$, we now know that for each $\lambda > 1$ there is a $\mathcal{W}$-null set $B(\lambda)$ such that $\overline{\lim_{m\to\infty}} \|\xi_{n_m(\lambda)}(\cdot,\omega) - K_T\|_T^0 = 0$ for each $\omega \notin B(\lambda)$. Set $B = \bigcup_1^\infty B(1 + 1/n)$. Then $\mathcal{W}(B) = 0$; and to each $\omega \notin B$ we can apply Lemma (1.20) with $\Lambda = \{1 + 1/n : n \ge 1\}$ and

thereby conclude both that the sequence $\{\xi_n(\cdot,\omega) : n \geq 2\}$ is precompact and that every convergent subsequence converges to an element of K_T.

To prove iii), we proceed as follows. First, note that there exists a countable, $\|\cdot\|_T^0$-dense subset of K_T consisting of ϕ's satisfying $2I_T(\phi) < 1$. (Indeed, simply take $\phi_n(t) = \int_0^{t\wedge T} \dot{\phi}_n(t)dt$, $n \geq 1$ and $t \geq 0$, where $\{\dot{\phi}_n\}_1^\infty \subseteq L^2([0,T];R^d)$ is dense in $\{\dot{\phi} \in L^2([0,T];R^d): \|\dot{\phi}\|_{L^2([0,T];R^d)} < 1\}$.) Thus, it suffices for us to show that for each $\phi \in K_T$ satisfying $2I_T(\phi) < 1$, $\mathfrak{W}(\underline{\lim}_{n\to\infty} \|\xi_n(\cdot) - \phi\|_T^0 = 0) = 1$.

Let $\phi \in K_T$ with $2I_T(\phi) < 1$ be given. For $k \geq 2$ define $\phi_k(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq T/k \\ \phi(t) - \phi(T/k) & \text{if } t \geq T/k \end{cases}$; and if $m \geq 1$ set $\eta_{m,k}(t) = \begin{cases} 0 & \text{if } 0 \leq t \leq T/k \\ (x(k^m t) - x(k^{m-1}T))/\phi(k^m) & \text{if } t \geq T/k \end{cases}$. Note that

$$\|\xi_{k^m}(\cdot,\omega) - \phi\|_T^0 \leq \|\xi_{k^m}(\cdot,\omega)\|_{T/k}^0 + \|\phi\|_{T/k}^0 + \|\xi_{k^m}(\cdot,\omega) - \phi\|_T^{T/k}$$

$$\leq 2(\|\xi_{k^m}(\cdot,\omega)\|_{T/k}^0 + \|\phi\|_{T/k}^0) + \|\eta_{m,k}(\cdot,\omega) - \phi_k\|_T^0 .$$

Thus if $C = \{\omega : \{\xi_n(\cdot,\omega)\}_2^\infty$ is $\|\cdot\|_T^0$-pre-compact$\}$, then for given $\delta > 0$ and any $\omega \in C$ we can choose $k(\omega) \leq 2$ such that

$$\|\xi_{k^m(\omega)}(\cdot,\omega) - \phi\|_T^0 \leq \delta + \|\eta_{m,k(\omega)} - \phi_k\|_T^0$$

for all $m \geq 1$. Since $\mathfrak{W}(C) = 1$, we now see that it suffices to prove that for each $k \geq 2$, $\mathfrak{W}(\underline{\lim}_{m\to\infty} \|\eta_{m,k}(\cdot) - \phi_k\|_T^0 = 0) = 1$. Noting that, for fixed $k \geq 2$, the processes $\{\eta_{m,k}(t) : 0 \leq t \leq T\}_{m=1}^\infty$ are independent under $\mathfrak{W}$, we will know that $\mathfrak{W}(\underline{\lim}_{m\to\infty} \|\eta_{m,k}(\cdot) - \phi_k\|_T^0 = 0) = 1$ once we show that for every $\varepsilon > 0$:

$$\sum_{m=1}^{\infty} \mathcal{W}(\|\eta_{m,k}(\cdot) - \phi_k\|_T^0 < \varepsilon) = \infty .$$

But

$$\mathcal{W}(\|\eta_{m,k}(\cdot) - \phi_k\|_T^0 < \varepsilon) = \mathcal{W}_{\varepsilon_m(k)}(\|x\ \ ) - \tilde{\phi}_k\|_{(1-1/k)T}^0 < \varepsilon) ,$$

where $\varepsilon_m(k) = (2 \log_2 k^m)^{-1}$ and $\tilde{\phi}_k(t) = \phi_k(t + T/k)$. Since $2I_{(1-1/k)T}(\tilde{\phi}_k) \le 2I_T(\phi) < 1$, we can use (1.6) to find $m(k)$ such that for $m \ge m(k)$:

$$\mathcal{W}(\|\eta_{m,k}(\cdot) - \tilde{\phi}_k\|_T^0 < \varepsilon) \ge e^{-\gamma \log_2 k^m} = \frac{1}{(m \log k)^\gamma}$$

where $\gamma \in (2I_{(1-1/k)T}(\phi),1)$. ■

Before dropping this topic, we will indicate how Strassen's theorem can be used to rederive the Hartman-Winter-Khinehin law of the iterated logarithm. The main new ingredient is the following theorem due to Skoroh d.

(1.22) Theorem: Let μ be a probability measure on R^1 satisfying $\int x^2\ \mu(dx) = 1$ and $\int x\ \mu(dx) = 0$. Then there is an R^1-valued Brownian motion on some probability space $(E,\mathcal{F},P)$ and there are independent identically distributed random times $\tau_n : E \to [0,\infty)$ such that $E^P[\tau_n] = 1$ and $\{\beta(\sum_1^n \tau_m) - \beta(\sum_1^{n-1} \tau_m)\}_1^n$ is a sequence of independent μ-distributed random variables on $(E,\mathcal{F},P)$.

Skorohod's proof of (1.22) consists of two parts. The first part is outlined in the next exercise.

(1.23) Exercise: Given $p,q \in [0,\infty)$, define $\nu_{p,q} = \frac{q}{p+q}\delta_p + \frac{p}{p+q}\delta_{-q}$ if $p \cdot q > 0$ and $\nu_{p,q} = \delta_0$ if $p \cdot q = 0$. Given a μ of the sort in theorem (1.22) , show that there exists a probability measure α on $[0,\infty) \times [0,\infty)$ such that $\mu = \int_{[0,\infty)^2} \nu_{p,q}\ \alpha(dp \times dq)$. The easiest way

to prove this representation is to first assume that μ is the sum of a finite number of atoms. One then uses weak convergence to extend it to μ's having compact support. Finally, using $\int x^2\,\mu(dx) = 1$, one extends it to the general case.

The second part of Skorokod's proof relies on the strong Markov property of Brownian motion: "a Brownian motion starts anew at a stopping time."

(1.24) Theorem: Let $(\beta(t),\mathcal{F}_t,P)$ be an R^d-valued Brownain motion. Given an $\mathcal{F}_t$-stopping time τ and an $A \in \mathcal{F}_\tau$ satisfying $A \subseteq \{\tau < \infty\}$ and $P(A) > 0$, define $Q^A = \chi_A\, P/P(A)$ $(\mathcal{F}_\tau = \{A \in \mathcal{F} : A \cap \{\tau \le t\} \in \mathcal{F}_t$ for all $t \ge 0\})$. Then $(\beta^\tau(t),\ \mathcal{F}_{t+\tau}, Q^A)$ is a Brownian motion, where $\beta^\tau(t) = \chi_{[0,\infty)}(\tau)(\beta(t+\tau) - \beta(\tau))$. In fact, if $A \in \mathcal{F}_\tau$ and $\Gamma \in \mathcal{M}$, then $P(A \cap \{\tau < \infty\} \cap \{\beta^\tau \in \Gamma\}) = P(A \cap \{\tau < \infty\})\cdot \mathcal{W}(\Gamma)$.

Proof: Suppose that we have proved the last part. Given $A \in \mathcal{F}_\tau$ such that $A \subseteq \{\tau < \infty\}$ and $P(A) > 0$, $0 \le t_1 < t_2$, and $B \in \mathcal{F}_{\tau+t_1}$ satisfying $P(A \cap B) > 0$, we would have, for all $\theta \in R^d$:

$$E^{Q^A}[e^{i(\theta,\beta^\tau(t_2))},B]$$

$$= \frac{P(A\cap B)}{P(A)} E^{Q^{A\cap B}}[e^{i(\theta,\beta^{\tau+t_1}(t_2-t_1))}\, e^{i(\theta,\beta^\tau(t_1))}]$$

$$= \frac{P(A\cap B)}{P(A)} E^{\mathcal{W}}[e^{i(\theta,x(t_2-t_1))}]\, E^{Q^{A\cap B}}[e^{i(\theta,\beta^\tau(t_1))}]$$

$$= \frac{P(A\cap B)}{P(A)} E^{Q^{A\cap B}}[\exp(i(\theta,\beta^\tau(t_1)) - \frac{|\theta|^2}{2}(t_2 - t_1))]$$

$$= E^{Q^A}[\exp(i(\theta,\beta^\tau(t_1)) - \frac{|\theta|^2}{2}(t_2 - t_1)),B]\,.$$

This is just the Fourier transform of b) in the definition of a Brownian motion; and therefore we would have proved that $(\beta^\tau(t),\mathcal{F}_{t+\tau},Q^A)$ is a Brownian motion. Thus, we need only prove the last part of the theorem.

In proving the last part of the theorem, we will assume that $\tau \le T$ for some $T > 0$, since in general we can replace τ by $\tau \wedge T$ and let $T \uparrow \infty$. Given $A \in \mathcal{F}_\tau$ with $P(A) > 0$, $0 \le t_1 < t_2$, and $\Gamma \in \mathcal{M}_{t_1}$, we have, by Doob's stopping time theorem:

$$E^{Q^A \circ \beta^\tau(\cdot)^{-1}}[e^{(\theta, x(t_2)) - \frac{|\theta|^2}{2} t_2}, \Gamma]$$

$$= E^{Q^A}[e^{(\theta, \beta^\tau(t_2)) - \frac{|\theta|^2}{2} t_2}, \beta^\tau(\cdot)^{-1}(\Gamma)]$$

$$= \frac{1}{P(A)} E^P[e^{(\theta, \beta(\tau + t_2) - \beta(\tau)) - \frac{|\theta|^2}{2}(\tau + t_2)} e^{\frac{|\theta|^2}{2}\tau}, A \cap \beta^\tau(\cdot)^{-1}(\Gamma)]$$

$$= \frac{1}{P(A)} E^P[e^{(\theta, \beta(\tau + t_1) - \beta(\tau)) - \frac{|\theta|^2}{2}(\tau + t_1)} e^{\frac{|\theta|^2}{2}\tau}, A \cap \beta^\tau(\cdot)^{-1}(\Gamma)]$$

$$= E^{Q^A}[e^{(\theta, \beta^\tau(t_1)) - \frac{|\theta|^2}{2} t_1}, \beta^\tau(\cdot)^{-1}(\Gamma)]$$

$$= E^{Q^A \circ \beta^\tau(\cdot)^{-1}}[e^{(\theta, x(t_1)) - \frac{|\theta|^2}{2} t_1}, \Gamma],$$

for all $\theta \in R^d$. Since $Q^A \circ \beta^\tau(\cdot)^{-1}(x(0) = 0) = 1$, this proves that $Q^A \circ \beta^\tau(\cdot)^{-1} = \mathcal{W}$. ■

(1.25) <u>Proof of</u> (<u>1.22</u>):

Take $d = 1$ in the definition of $(\Omega, \mathcal{M}, \mathcal{W})$. Set $E = \Omega \times ([0,\infty)^2)^{Z^+}$, $\mathcal{F} = \mathcal{M} \times \mathcal{B}_{([0,\infty)^2)^{Z^+}}$, and $\mathcal{F}_t = \mathcal{M}_t \times \mathcal{B}_{([0,\infty)^2)^{Z^+}}$, $t \ge 0$..

Given μ, choose α on $[0,\infty)^2$ as in (1.23), and set $P = \mathcal{W} \times \alpha^{Z^+}$. Then $(\beta(t), \mathcal{F}_t, P)$ is a Brownian motion on $(E, \mathcal{F}\, P)$, where $\beta(\cdot, \omega, (\vec{p}, \vec{q})) = x(\cdot, \omega)$. Next, set $T_0 \equiv 0$ and

$$T_n(\omega, (\vec{p}, \vec{q}))$$

$$= \inf\{t \ge T_{n-1}(\omega, (\vec{p}, \vec{q})) : x(t, \omega) - x(T_{n-1}(\omega, (\vec{p}, \vec{q})), \omega) \notin (-q_n, p_n)\}.$$

Clearly the T_n's are $\mathcal{F}_t$-stopping times. Moreover, if $\tau_n = T_n - T_{n-1}$,

then $\tau_1(\omega,(\vec{p},\vec{q}))$ is the first exit time of $x(\cdot,\omega)$ from $(-q_1,p_1)$, and therefore

$$\int \tau_1(\omega,(\vec{p},\vec{q}))\mathcal{W}(d\omega) = p_1\, q_1 = \int x^2\, \nu_{p_1,q_1}(dx)\ .$$

Hence, $E^P[\tau_1] = \int x^2\, \mu(dx) = 1$. Assuming that $T_n < \infty$ (a.s., P), one can use (1.24) to conclude that τ_{n+1} and $\beta(\cdot + T_n) - \beta(T_n)$ are independent of $\mathcal{F}_{T_n}$ and have the same distribution as τ_1 and $\beta(\tau_1)$, respectively. Since $x(\tau_1(\cdot,(\vec{p},\vec{q})),\cdot)$ has distribution ν_{p_1,q_1} and therefore $\beta(\tau_1)$ has distribution μ, we now see that the $\beta(T_n) - \beta(T_{n-1})$'s are independent and have distribution μ. ∎

We can now prove the following important corollary of Strassen's basic result.

(1.26) Theorem: Let $X_1,\ldots,X_n,\ldots$ be independent, identically distributed, R^1-valued random variables having mean 0 and variance 1. Define $S(t)$, $t \geq 0$, by

$$S(t) = (1 - t - [t])S_{[t]} + (t - [t])S_{[t]+1}$$

where $S_n = \sum_1^n X_k$ and $S_0 \equiv 0$. For $T > 0$ and $n \geq 2$, set

$$\eta_n(t) = S(nt)/\phi(n)\ ,\quad 0 \leq t \leq T$$

where $\phi(n) = (2n \log_2 n)^{1/2}$. Then, almost surely, one has

i) $\{\eta_n(\cdot) : n \geq 2\}$ is pre-compact in $C([0,T];R^1)$;

ii) any convergent subsequence of $\{\eta_n(\cdot): n \geq 2\}$ converges to to an element of K_T;

iii) for every $\psi \in K_T$, there is a subsequence of $\{\eta_n(\cdot): n \geq 2\}$ which converges to ψ.

(The set K_T is the same as the one in Theorem (1.17).) In particular, if $\Phi : C([0,T] ; R^1) \to R^1$ is continuous, then $P(\lim_{n\to\infty} \Phi(\eta_n(\cdot)) = \sup_{\Phi \in K_T} \Phi(\phi)) = 1$.

Proof: In view of Theorem (1.22), we can find $(\beta(t), \mathcal{F}_t, P)$ and $\{\tau_n\}_1^\infty$ as in that theorem so that $\{S(n)\}_1^\infty$ and $\{\beta(\sum_1^n \tau_k)\}_1^\infty$ have the same distribution. Thus we will assume that $S(n) = \beta(\sum_1^n \tau_k)$. It is clear that the proof will be complete if we show that $\overline{\lim_{n\to\infty}} \sup_{0\le t\le T} \frac{|S(nt) - \beta(nt)|}{\varphi(n)} = 0$ (a.s., P) . For notational convenience, we will do this only when $T = 1$.

Note that:

$$\sup_{0\le t\le 1} |S(nt) - \beta(nt)| = \max_{1\le m\le n} \sup_{m-1\le t\le m} |S(t) - \beta(t)|$$

$$\le \max_{1\le m\le n} |S(m) - \beta(m)| \vee |S(m-1) - \beta(m)|$$

$$+ \max_{1\le m\le n} \sup_{m-1\le t\le m} |\beta(t) - \beta(m)|$$

Thus:

$$\sup_{0\le t\le 1} \frac{|S(nt) - \beta(nt)|}{\varphi(n)}$$

$$\le \max_{1\le m\le n} |\xi_n(\tfrac{1}{n}\sum_1^m \tau_k) - \xi_n(\tfrac{m}{n})| \vee |\xi_n(\tfrac{1}{n}\sum_1^{m-1} \tau_k) - \xi_n(\tfrac{m}{n})|$$

$$+ \max_{1\le m\le n} \sup_{m-1\le t\le m} |\xi_n(t/n) - \xi_n(m/n)|$$

$$\le \max_{1\le m\le n} |\xi_n(\tfrac{1}{n}\sum_1^m \tau_k) - \xi_n(\tfrac{m}{n})|$$

$$+ 2 \sup_{\substack{0\le s\le t\le 1 \\ t-s\le 1/n}} |\xi_n(s) - \xi_n(t)| ,$$

where $\xi_n(t) = \beta(nt)/\varphi(n)$, $0 \le t \le 1$. Since, P-almost surely, $\{\xi_n(\cdot)\}_2^\infty$ is pre-compact in $C([0,1] ; R^1)$, we see that

$$\overline{\lim_{n\to\infty}} \sup_{0\le t\le 1} \frac{|S(nt) - \beta(nt)|}{\varphi(n)} \le \overline{\lim_{n\to\infty}} \max_{1\le m\le n} |\xi_n(\frac{1}{n}\sum_1^m \tau_k) - \xi_n(\frac{m}{n})|$$

(a.s., P) . Finally, given $\varepsilon > 0$, define $A_\varepsilon(N,\delta)$, for $N \ge 2$ and $\delta > 0$, to be the set $\{\sup_{m\ge N} |\frac{1}{m}\sum_1^m \tau_k - 1| < \delta\} \cap \{|\xi_n(t) - \xi_n(s)| \le \varepsilon$ for $n \ge N$ and $0 \le s < t \le 1$ satisfying $t - s < \delta\}$. By Theorem (1.17) and the strong law of large numbers, $P(A_\varepsilon(N,\delta)) \to 1$ as $N \uparrow \infty$ and $\delta \downarrow 0$. At the same time, $\{\overline{\lim_{n\to\infty}} \max_{1\le m\le n} |\xi_n(\frac{1}{n}\sum_1^m \tau_k) - \xi_n(\frac{m}{n})| \le \varepsilon\} \supseteq A_\varepsilon(N,\delta)$ (a.s., P) for each N and $\delta > 0$. □

(<u>1.27</u>) <u>Exercise</u>: Let $X_1,\ldots,X_n,\ldots$ be as in Theorem (1.26) and define $S(t)$, $t \ge 0$, accordingly. For $n \ge 1$, set $Y_n(t) = S(nt)/n^{1/2}$, $t \ge 0$, and denote by P_n the distribution of $Y_n(\cdot)$ on $(\Omega,\mathcal{M})$. Using the representation technique introduced in the proof of (1.26) , show that P_n tends weakly to $\mathbb{W}$ as $n \to \infty$. This result is known as Donsker's <u>invariance</u> <u>principle</u>.

2. Large Deviations, Some Generalities:

Let X be a Polish space with Borel field $\mathcal{B}$. We will say that the function $I : X \to [0,\infty]$ is a <u>rate function</u> if

i) $I \not\equiv \infty$,

ii) I is lower semi-continuous,

iii) for any $L \geq 0$, $\{x : I(x) \leq L\}$ is compact.

A family $\{\mu_\varepsilon : \varepsilon > 0\}$ of probability measures on $(X,\mathcal{B})$ is said to satisfy the <u>large deviations principle with rate I</u> if

$$\overline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log \mu_\varepsilon (F) \leq - \inf_{x\in F} I(x) \tag{2.1}$$

for all closed sets F in X , and

$$\underline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log \mu_\varepsilon (G) \geq - \inf_{x\in G} I(x) \tag{2.2}$$

for all non-empty open G in X .

<u>(2.3)</u> <u>Remark:</u> As we have seen in section 1) , it is often convenient not to restrict one's attention to X but, instead, to a larger space. For instance, in our discussion of Schilder's Theorem, we thought of the $\mathcal{W}_\varepsilon$'s as living on $\Omega = C([0,\infty); R^d)$. However, Schilder's theorem really involves the restriction of the $\mathcal{W}_\varepsilon$'s to $X_T = \{\psi \in C([0,T] ; R^d) : \psi(0) = 0\}$ for some $T > 0$ and says that the $\mathcal{W}_\varepsilon|_{X_T}$'s satisfy the large deviations principle with rate I_T.

<u>(2.4)</u> <u>Remark:</u> We have already seen how important conditions ii) and iii) are. For example, by precisely the same argument as we used to derive (1.10) :

$$\inf_{x\in F^\delta} I(x) \uparrow \inf_{x\in F} I(x) \quad \text{as} \quad \delta \downarrow 0 \tag{2.5}$$

for any closed $F \neq \emptyset$ in X .

We next show how our set-up leads to an interesting abstraction of Laplace's asymptotics.

(2.6) Theorem. (Varadhan): Let $\{\mu_\varepsilon : \varepsilon > 0\}$ satisfy the large deviation principle with rate I . If $\Phi : X \to R^1$ is a bounded continuous function, then

$$\lim_{\varepsilon\downarrow 0} \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] = \sup_{x\in X}(\Phi(x) - I(x)) \quad . \tag{2.7}$$

Proof: Set $\ell = \sup_{x\in X} (\Phi(x) - I(x))$ and define $K = \{x : I(x) \leq \|\Phi\| + \ell\}$. Then K is compact. Given $\delta > 0$, cover K with a finite number of open sets $G_1,\ldots,G_N$ such that $\sup_{x,y\in G_n} |\Phi(x) - \Phi(y)| \leq \delta$ for all $1 \leq n \leq N$; set $F = (\bigcup_1^N G_n)^c$; and for $1 \leq n \leq N$, define $a_n = \inf_{x\in\overline{G}_n} \Phi(x)$ and $b_n = \inf_{x\in\overline{G}_n} I(x)$. Choose $\varepsilon_o > 0$ so that

$$\mu_\varepsilon(\overline{G}_n) \leq e^{-(b_n-\delta)/\varepsilon}$$

and

$$\mu_\varepsilon(F) \leq \exp[- (\inf_{x\in F} I(x)-\delta)/\varepsilon]$$

for $0 < \varepsilon < \varepsilon_o$. Since $\inf_{x\in F} I(x) \geq \|\Phi\| + |\ell|$, we have:

$$E^{\mu_\varepsilon}[\exp (\Phi/\varepsilon)] \leq \sum_1^N E^{\mu_\varepsilon}[\exp (\Phi/\varepsilon), \overline{G}_n] + E^{\mu_\varepsilon}[\exp (\Phi/\varepsilon), F]$$

$$\leq \sum_{1}^{N} e^{(a_n+2\delta-b_n)/\varepsilon} + e^{(\delta-|\ell|)/\varepsilon}$$

$$= e^{(\ell+2\delta)/\varepsilon}\left[\sum_{1}^{N} e^{(a_n-b_n-\ell)/\varepsilon} + e^{-(\delta+|\ell|+\ell)/\varepsilon}\right]$$

for $0 < \varepsilon \leq \varepsilon_o$. Note that $a_n - b_n \leq \sup_{x\in G_n}(\Phi(x)-I(x)) \leq \ell$, and so this proves that

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log\left(E^{\mu_\varepsilon}\left[e^{\Phi/\varepsilon}\right]\right) \leq \ell+2\delta.$$

Since this is true for all $\delta > 0$, we have:

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log\left(E^{\mu_\varepsilon}\left[e^{\Phi/\varepsilon}\right]\right) \leq \ell \quad .$$

Next, given $\delta > 0$, choose $x_o \in X$ so that $\Phi(x_o) - I(x_o) \geq \ell - \delta$. Now, choose an open set $U \ni x_o$ so that $|\Phi(x) - \Phi(x_o)| < \delta$ for all $x \in U$. Then:

$$\underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log E^{\mu_\varepsilon}\left[e^{\Phi/\varepsilon}\right] \geq \underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log E^{\mu_\varepsilon}\left[e^{\Phi/\varepsilon},U\right]$$

$$\geq \Phi(x_o) - \delta + \underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log \mu_\varepsilon(U)$$

$$\geq \Phi(x_o) - \delta - \inf_{x\in U} I(x)$$

$$\geq \Phi(x_o) - I(x_o) - \delta$$

$$\geq \ell - 2\delta \quad . \quad \blacksquare$$

Before proving a useful corollary of (2.6), we need the following.

(2.8) <u>Lemma</u>: Let $\Phi : x \to R^1\cup\{+\infty\}$ be a function which is bounded below. Then Φ is lower semi-continuous if and only if there exists $\{\Phi_n\}_1^\infty \subseteq C_b(X)$ such that $\Phi_n\uparrow\Phi$.

<u>Proof</u>: Say there exists $\{\Phi_n\}_1^\infty \subseteq C_b(x)$ such that $\Phi_n \uparrow \Phi$. Then $\{\Phi \leq c\} = \bigcap_n\{\Phi_n \leq c\}$ is closed for each $c \in R^1$; that is, Φ is ℓ.s.c.

To prove the converse, first suppose that $\Phi = \chi_G$ for some open G in X. For $n \geq 1$, set $\Phi_n(x) = \dfrac{\operatorname{dist}(x,G^c)}{(1/n) \vee \operatorname{dist}(x,G^c)}$, $x \in X$. Clearly $\Phi_n \in C_b(X)$ and $\Phi_n \uparrow \Phi$. Next, let Φ be any ℓ.s.c. function which is bounded below. Since $\Phi \wedge n \uparrow \Phi$ as $n \to \infty$, and $\Phi \wedge n$ is ℓ.s.c. for each $n \geq 1$, assume that Φ is bounded. In fact, without loss of generality, we will assume that $\Phi(x) \in (0,1)$, $x \in X$. Given $n \geq 1$, set

$$\Phi_n = \sum_{k=0}^{n-1} \frac{k}{n} \chi_{(\frac{k}{n}, \frac{k+1}{n}]}(\Phi) .$$

Clearly $\Phi_n \uparrow \Phi$. Moreover, we can re-write Φ_n as $\frac{1}{n} \sum_{k=1}^{n} \chi_{(\frac{k}{n},1)}(\Phi)$. Since $\Phi(x) \in (0,1)$, $x \in X$, and Φ is ℓ.s.c., $\{x : \Phi(x) \in (\frac{k}{n},1)\} = \{x : \Phi(x) > \frac{k}{n}\}$ is open. Hence each Φ_n is the increasing limit of a sequence from $C_b(X)$. ∎

(2.9) Corollary: Let $\{\mu_\varepsilon : \varepsilon > 0\}$ and I be as in Theorem (2.6). If $\Phi : X \to R^1$ is lower semi-continuous and bounded below, then:

$$\underline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] \geq \sup_{\substack{x \in X \\ \Phi(x) \wedge I(x) < \infty}} (\Phi(x) - I(x)) . \tag{2.10}$$

(We adopt the convention that the supremum over $\emptyset$ is $-\infty$.)

If $\Phi : X \to R^1$ is upper semi-continuous and bounded above, then

$$\overline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] \leq \sup_{x \in X} (\Phi(x) - I(x)) . \tag{2.11}$$

Proof: In the first case, choose $\{\Phi_n\}_1^\infty \subseteq C_b(x)$ so that $\Phi_n \uparrow \Phi$. Then

$$\underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] \geq \sup_{x\in X} (\Phi_n(x) - I(x)) \quad .$$

Assuming that $\ell \equiv \sup_{\substack{x\in X \\ \Phi(x)\wedge I(x)<\infty}} (\Phi(x) - I(x)) > -\infty$, we have that $\ell = \sup_{\substack{x\in X \\ I(x)<\infty}} (\Phi(x) - I(x))$. Hence for $\delta > 0$, we can find $x_\delta \in X$ so that $I(x_\delta) < \infty$ and:

$$\Phi(x_\delta) \geq \begin{cases} 1/\delta + I(x_\delta) & \text{if } \ell = \infty \\ \ell - \delta + I(x_\delta) & \text{if } \ell < \infty \quad . \end{cases}$$

Thus, we can find $n \geq 1$ so that

$$\Phi_n(x_\delta) \geq \begin{cases} 1/2\delta + I(x_\delta) & \text{if } \ell = \infty \\ \ell - 2\delta + I(x_\delta) & \text{if } \ell < \infty \quad . \end{cases}$$

We therefore see that so long as $\ell > -\infty$, (2.10) holds. But, if $\ell = -\infty$, then (2.10) is trivial.

We now turn to the second case. Set $\ell = \sup_{x\in X} (\Phi(x) - I(x))$. If $\ell = -\infty$, then, for any $L \geq 0$, $\{x : \Phi(x) \geq -L\} \subseteq \{x : I(x) = +\infty\}$. Since Φ is u.s.c., $\{x : \Phi(x) \geq -L\}$ is closed. Thus

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}, \Phi \geq -L] = -\infty$$

for all L . Therefore, for each $L \geq 0$ there is an $\varepsilon_L > 0$ such that

$$E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] = E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}, \Phi < -L] + E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}, \Phi \geq -L] \leq 2e^{-L/\varepsilon}$$

if $0 < \varepsilon \leq \varepsilon_L$. Hence $\overline{\lim\limits_{\varepsilon \downarrow 0}}\ \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] = -\infty$ if $\ell = -\infty$. Now assume that $\ell > -\infty$. Then $\infty > \ell = \sup\{\Phi(x) - I(x) : \Phi(x) > -\infty$ and $I(x) < \infty\}$. Choose $\{\Phi_n\}_1^\infty \subseteq C_b(X)$ so that $\Phi_n \downarrow \Phi$ and note that

$$\overline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] \leq \sup_{x \in X} (\Phi_n(x) - I(x))$$

for each $n \geq 1$. Given $n \geq 1$, choose x_n so that $I(x_n) < \infty$ and $\Phi_n(x_n) - I(x_n) \geq \sup\limits_{x \in X} (\Phi_n(x) - I(x)) - 1/n$. Since $\sup\limits_{x \in X} (\Phi_n(x) - I(x)) \geq \ell$, $I(x_n) \leq M - \ell$ where $M = \sup\limits_{x \in X} \Phi_0(x) < \infty$. Because $\ell > -\infty$, this shows that $\{x_n\}_1^\infty$ is pre-compact. Choose a subsequence $\{x_{n'}\}$ which converges to some x_0 . Then, since $-I$ is upper semi-continuous,

$$\overline{\lim_{n' \to \infty}}\ [\Phi_{n'}(x_{n'}) - I(x_{n'})] \leq \Phi_m(x_0) - I(x_0)$$

for every $m \geq 1$. Hence:

$$\begin{aligned} \overline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] &\leq \overline{\lim_{n' \to \infty}}\ (\Phi_{n'}(x_{n'}) - I(x_{n'})) \\ &\leq \Phi_m(x_0) - I(x_0) \end{aligned}$$

for all $m \geq 1$; and so

$$\overline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log E^{\mu_\varepsilon}[e^{\Phi/\varepsilon}] = \Phi(x_0) - I(x_0) \leq \ell \quad . \blacksquare$$

(2.12) Remark: If Φ is upper semi-continuous and bounded above, then there is an $x_0 \in X$ such that $\Phi(x_0) - I(x_0) = \sup\limits_{x \in X} (\Phi(x) - I(x))$. Indeed, if $\sup\limits_{x \in X} (\Phi(x) - I(x)) = -\infty$, then $\Phi(x) - I(x) = -\infty$ for every $x \in X$. If $\sup\limits_{x \in X} (\Phi(x) - I(x)) = \ell > -\infty$, we can find $\{x_n\}_1^\infty \subseteq X$ so that $I(x_n) < \infty$

and $\Phi(x_n) - I(x_n) \geq \ell - 1/n$. Since Φ is bounded above by some $M < \infty$, $\{x_n\}_1^\infty \subseteq \{x : I(x) \leq M - \ell + 1\}$ and so $\{x_n\}_1^\infty$ is pre-compact. Choose a convergent subsequence $\{x_{n'}\}$ with limit x_0 . By the upper semi-continuity of $\Phi - I$, $\Phi(x_0) - I(x_0) \geq \overline{\lim_{n'\to\infty}} (\Phi(x_{n'}) - I(x_{n'})) \geq \ell$.

3. Cramér's Theorem:

There are two directions in which one might try to extend the results obtained thus far. One of these is to see how far one can go on reasoning based on independent increments. The second is to generalize the results of section 1) by taking advantage of Gaussian properties. In this section, we will go in the first of these directions.

We begin with the following. Let μ be a probability measure on R^1 with the property that

$$M(\xi) \equiv \int e^{\xi x}\mu(dx) < \infty$$

for all $\xi \in R^1$. Note that

(3.1)

i) $M \in C^{\infty}(R^1)$ and has values in $(0,\infty)$,

ii) $M(\alpha\xi_1 + (1-\alpha)\xi_2) \leq M(\xi_1)^{\alpha}M(\xi_2)^{1-\alpha}$ for all $\xi_1,\xi_2 \in R^1$ and $0 < \alpha < 1$.

The first assertion in (3.1) is an easy consequence of Lebesgue's Dominated Convergence Theorem, the second is a simple application of Hölder's inequality. We next define

$$I(x) = \sup_{\xi\in R^1}(\xi x - \log M(\xi)) \quad , \quad x \in R^1 . \tag{3.2}$$

(3.3) Lemma: The function I is a rate function. Moreover, I is convex. Finally, if $a = \int x\mu(dx)$, then $I(a) = 0$; and therefore I is non-decreasing on $[a,\infty)$ and non-increasing on $(-\infty,a]$.

Proof: Since $\xi x - \log M(\xi) = 0$ when $\xi = 0$ no matter what x is, $I \geq 0$. Also, as the supremum of continuous functions, I is l.s.c. Now suppose that $L \in [0,\infty)$ and consider $K_L = \{x : I(x) \leq L\}$. Since I is l.s.c, K_L is closed. Moreover, $I(x) \leq L$ implies that $x - \log M(1) \leq L$ and that $-x - \log M(-1) \leq L$. Thus, $K_L \subseteq [-L - \log M(-1), L + \log M(1)]$, and so K_L is bounded. We have therefore shown that I is a rate function. To see that I is convex, let $x_1, x_2 \in R^1$ and $0 < \alpha < 1$ be given. Then, for any $\xi \in R^1$:

$$\begin{aligned}\alpha I(x_1) + (1-\alpha)I(x_2) &\geq \alpha(\xi x_1 - \log M(\xi)) + (1-\alpha)(\xi x_2 - \log M(\xi)) \\ &= \xi(\alpha x_1 + (1-\alpha)x_2) - \log M(\xi)) ,\end{aligned}$$

and so $I(\alpha x_1 + (1-\alpha)x_2) \leq \alpha I(x_1) + (1-\alpha)I(x_2)$.

Finally, by Jensen's inequality, $\log(\int e^{\xi x}\mu(dx)) \geq \xi\int x\mu(dx) = \xi a$. Thus, $\xi a - \log M(\xi) \leq 0$ for all $\xi \in R^1$, and so $I(a) = 0$. Also, if $a < x_1 < x_2$, choose $0 < \alpha < 1$ so that $x_1 = (1-\alpha)a + \alpha x_2$. Then $I(x_1) \leq (1-\alpha)I(a) + \alpha I(x_2) = \alpha I(x_2) \leq I(x_2)$. Similarly, if $x_1 < x_2 < a$, then $I(x_1) \geq I(x_2)$. ∎

Next, let $X_1,\ldots,X_n$ be independent random variables, each with distribution μ. Set $S_n = \sum_1^n X_m$ and let μ_n be the distribution of S_n/n. We are going to show that $\{\mu_n : n \geq 1\}$ satisfies the large deviation principle with rate I (take $\varepsilon = 1/n$).

(3.4) <u>Lemma</u>: Let F be a closed subset of R^1. Then

(3.5) $$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} I(x) \quad .$$

<u>Proof</u>: If $F = \emptyset$ or $a \in F$, there is nothing to prove. Assume that $F \neq \emptyset$ and that $a \notin F$. Assume, for the moment, that $F \subseteq (a,\infty)$ and set $y_2 = \inf\{y : y \in F\}$. Then, for $\xi \geq 0$:

$$\mu_n(F) \leq \mu_n([y_2,\infty)) \leq e^{-\xi y_2} E[e^{\xi S_n/n}]$$

$$= e^{-\xi y_2} M(\xi/n)^n$$

and so

$$\frac{1}{n} \log \mu_n(F) \leq -(\xi/n\, y_2 - \log M(\xi/n))$$

for all $\xi \geq 0$. Hence

$$\frac{1}{n} \log \mu_n(F) \leq -\sup_{\xi\geq 0}(\xi y_2 - \log M(\xi)) \quad .$$

Since $y_2 > a$ and therefore

$$\xi y_2 - \log M(\xi) \leq \xi y_2 - \xi a = \xi(y_2-a) < 0$$

if $\xi < 0$, $I(y_2) = \sup_{\xi\geq 0}(\xi y_2 - \log M(\xi))$. Thus, we have

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -I(y_2) \quad .$$

By a similar argument, if $F \subseteq (-\infty,a)$, then

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -I(y_1) \quad ,$$

where $y_1 = \sup\{y : y \in F\}$. By the last part of Lemma (3.3) , this finishes the proof when $F \subseteq (a,\infty)$ or $F \subseteq (-\infty,a)$.

To complete the proof, let $F \neq \emptyset$ be a closed set not containing a such that neither $F_1 = (-\infty,a) \cap F$ nor $F_2 = F \cap (a,\infty)$ is empty. Let $y_1 = \sup\{y : y \in F_1\}$ and $y_2 = \inf\{y : y \in F_2\}$.

By the preceding paragraph:

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log \left[\mu_n(F_1) \vee \mu_n(F_n)\right]$$

$$\leq -(I(y_1) \wedge I(y_2)) .$$

Finally, note that, by the last part of Lemma (3.3) , $I(y_1) = \inf_{x\in F_1} I(x)$ and and $I(y_2) = \inf_{x\in F_2} I(x)$. ■

(3.6) <u>Lemma</u>: Let $G \neq \emptyset$ be open. Then

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \geq -\inf_{x\in G} I(x) \tag{3.7}$$

<u>Proof</u>: We will show that for each $x \in G$, $\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \geq -I(x)$. If $I(x) = \infty$, there is nothing to do. Thus we assume that $I(x) < \infty$.

First, assume that there is no $\xi \in R^1$ such that $I(x) = \xi x - \log M(\xi)$. Then $x \neq a$, since $I(a) = 0 = \xi a - \log M(\xi)$ when $\xi = 0$. Assume that $x > a$. Then, there exists a sequence $\{\xi_n\}_1^\infty \in (0,\infty)$ tending to $+\infty$ such that $\xi_n x - \log M(\xi_n) \to I(x)$. Note that

$$\int_{(-\infty,x)} e^{\xi_n(y-x)} \mu(dy) \to 0 .$$

Also,

$$\lim_{n\to\infty} \int_{[x,\infty)} e^{\xi_n(y-x)} \mu(dy) = e^{-I(x)} < \infty$$

for all n . Thus $\mu((x,\infty)) = 0$; and so

$$e^{-I(x)} = \lim_{n\to\infty} \int e^{\xi_n(y-x)} \mu(dy) = \mu(\{x\}) \quad .$$

But this means that

$$\mu_n(G) \geq \mu_n(\{x\}) \geq \mu(\{x\})^n = e^{-nI(x)} \quad ,$$

which certainly implies that $\underline{\lim}_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq -I(x)$. A similar argument applies to $x < a$.

We now turn to the case when there is a $\xi_0 \in R^1$ such that $\xi_0 x - \log M(\xi_0) = I(x)$. Note that $x = M'(\xi_0)/M(\xi_0)$. Set $\tilde{\mu}(dy) = \frac{e^{\xi_0 y}}{M(\xi_0)} \mu(dy)$, and note that $\int y\tilde{\mu}(dy) = \frac{M'(\xi_0)}{M(\xi_0)} = x$. Assuming that $x \geq a$, and therefore that $\xi_0 \geq 0$, we have, for all sufficiently small $\delta > 0$:

$$\mu_n(G) \geq \mu_n((x-\delta,x+\delta)) = \int_{\{|\frac{1}{n}\sum_1^n y_m - x|<\delta\}} \mu(dy_1)\ldots\mu(dy_n)$$

$$\geq e^{-n\xi_0(x+\delta)} \int_{\{|\frac{1}{n}\sum_1^n y_m - x|<\delta\}} e^{\xi_0 y_1}\cdots e^{\xi_0 y_n} \mu(dy_1)\cdots\mu(dy_n)$$

$$= e^{-n\xi_0(x+\delta)} M(\xi_0)^n \int_{\{|\frac{1}{n}\sum_1^n y_m - x|<\delta\}} \tilde{\mu}(dy_1)\cdots\tilde{\mu}(dy_n) \quad .$$

By the law of large numbers,

$$\int_{\{|\frac{1}{n}\sum_1^n y_m - x|<\delta\}} \tilde{\mu}(dy_1)\cdots\tilde{\mu}(dy_n) \to 1$$

as $n \to \infty$. Hence

$$\varliminf_{n\to\infty} \frac{1}{n} \log \mu_n(G) \geq -\xi_0(x+\delta) + M(\xi_0) = -I(x) - \xi_0\delta$$

for every $\delta > 0$. A similar argument works if $x < a$ (and therefore $\xi_0 \leq 0$). ∎

We have now proved the following theorem.

(3.8) Theorem (Cramér): Let μ be a probability measure on R^1 satisfying $M_\mu(\xi) \equiv \int e^{\xi y}\mu(dy) < \infty$ for all $\xi \in R^1$. For $x \in R^1$, define

$$I_\mu(x) = \sup_{\xi\in R^1} (\xi x - \log M_\mu(\xi)) \quad .$$

Then I_μ is a rate function. Moreover, if μ_n denotes the distribution of $\frac{1}{n}\sum_1^n y_m$ under μ^n , then μ_n satisfies the large deviation principle with rate I_μ (take $\varepsilon = 1/n$).

(3.9) Exercise: Use Theorem (3.8) to prove the upper bound in the classical law of the iterated logarithm for independent, identically distributed R^1-valued random variables $X_1,\ldots,X_n,\ldots$ having common distribution μ , where $\int e^{\xi y}\mu(dy) < \infty$ for all $\xi \in R^1$, $\int y\mu(dy) = 0$, and $\int y^2\mu(dy) = 1$. That is, show that $\varlimsup_{n\to\infty} \sum_1^n X_m/(2n\log_2 n)^{1/2} \leq 1$ almost surely.

(3.10) Exercise:

i) When $\mu = p\delta_a + (1-p)\delta_b$, with $a < b$ and $p \in (0,1)$, show that

$$I\mu(x) = \begin{cases} \frac{x-a}{b-a} \log\cdot\frac{x-a}{1-p} + \frac{b-x}{b-a} \log \frac{b-x}{p} - \log(b-a) & \text{if } x \in (a,b) \\ -\log p & \text{if } x = a \\ -\log(1-p) & \text{if } x = b \\ \infty & \text{if } x \notin [a,b] \end{cases} .$$

ii) When $\mu(dx) = \chi_{[0,\infty)}(x)e^{-x}dx$, show that

$$I_\mu(x) = \begin{cases} x - 1 - \log x & , \quad x > 0 \\ \infty & , \quad x \le 0 \end{cases} .$$

iii) If $\mu(dx) = \frac{1}{(2\pi\sigma^2)^{1/2}} \exp(-(x-a)^2/2\sigma^2)dx$, show that

$$I_\mu(x) = \frac{1}{2\sigma^2}(x-a)^2 .$$

iv) In general, show that $I_{\delta_a^*\mu}(x) = I_\mu(x-a)$ for any $a \in R^1$

Our next goal is to extend Cramér's theorem to Banach space valued random variables. The procedure we will adopt is due to R.R. Bahadur and S.L. Zabell. The idea is to first show that $\{\mu_n : n \ge 1\}$ satisfies a large deviation principle with respect to some rate function and then identify the rate function.

(3.11) Lemma: Let $f : Z^+ \to R^1 \cup \{\infty\}$ be sub-additive (i.e. $f(m+n) \le f(m) + f(n)$ for all $m,n \in Z^+$). If for some $n_0 \in Z^+$, $f(n) < \infty$ for all $n \ge n_0$, $\lim_{n\to\infty} \frac{f(n)}{n}$ exists.

Proof: Let $\beta = \inf\{\frac{f(n)}{n} : n \ge n_0\}$. Given $m \ge n_0$, define $q_n = [n/m]$ and $r_n = n - [n/m]m$ for $n \ge m$. Then for $n \ge 2m$

$$\frac{f(n)}{n} = \frac{f(q_n m+r_n)}{n} \le \frac{f((q_n-1)m)+f(m+r_n)}{n}$$

$$\leq \frac{(q_n-1)m}{n}\frac{f(m)}{m} + \max_{m\leq \ell<2m}\frac{f(\ell)}{n} \quad ;$$

and so

$$\overline{\lim_{n\to\infty}}\frac{f(n)}{n} \leq \frac{f(m)}{m}$$

for all $m \geq n_0$. Hence:

$$\beta \leq \lim_{n\to\infty}\frac{f(n)}{n} \leq \overline{\lim_{n\to\infty}}\frac{f(n)}{n} \leq \inf_{m\geq n_0}\frac{f(m)}{m} = \beta \quad . \blacksquare$$

We remark that the condition $f(n) < \infty$ for sufficiently large n's is vital. Indeed, consider $f(n) = \begin{cases}\infty & \text{if } n \text{ is odd}\\ 0 & \text{if } n \text{ is even}\end{cases}$.

Let $(\Omega,\mathcal{M},P)$ be a probability space and let $(E,\|\cdot\|)$ be a separable Banach space. Given independent, identically distributed E-valued random variables $X_1,\ldots,X_n,\ldots$ on $(\Omega,\mathcal{M},P)$, set $\overline{X}_n = \frac{1}{n}\sum_1^n X_\ell$. The following lemma is pivotal.

(3.12) Lemma: Given a measurable convex set A in E , set $f_A(n) = -\log P(\overline{X}_n \in A)$. Then f_A is subadditive. Thus, if either $f_A \equiv \infty$ or $f_A(n) < \infty$ for sufficiently large n's , then $\lim_{n\to\infty}\frac{1}{n}\log P(\overline{X}_n \in A)$ exists. In particular, $\lim_{n\to\infty}\frac{1}{n}\log P(X_n \in A)$ exists for all open convex A's.

Proof: Note that $\overline{X}_{m+n} = \frac{m}{m+n}\overline{X}_m + \frac{n}{m+n}\overline{X}^m_{m+n}$ where $\overline{X}^m_{m+n} = \frac{1}{n}\sum_{k=1}^n X_{m+k}$; and therefore, since A is convex, $\{\overline{X}_m \in A\} \cap \{\overline{X}^m_{m+n} \in A\} \subseteq \{\overline{X}_{m+n} \in A\}$. But $P(\overline{X}_m \in A,\overline{X}^m_{m+n} \in A) = P(\overline{X}_m \in A)P(\overline{X}_n \in A)$, and so $P(\overline{X}_m \in A)P(\overline{X}_n \in A) \leq$ $P(\overline{X}_{m+n} \in A)$. This proves that f_A is subadditive. In view of Lemma (3.11) , it remains only to prove that if A is open and convex, then either $f_A \equiv \infty$ or $f_A(n) < \infty$ for sufficiently large n's . To this end, suppose

that $P(\overline{X}_m \in A) > 0$ for some m. Then $A \neq \emptyset$ and we can find a bounded closed convex set F in A such that $\mathrm{dist}(F,A^c) > 0$ and $P(\overline{X}_m \in F) > 0$. Set $B = \sup_{x\in F}\|x\|$ and choose $\delta > 0$ so that $\mathrm{dist}(F,A^c) > 2\delta$. Next, select

$n_0 > m$ so that $\alpha \equiv \min_{1\leq \ell\leq m} P(\|\frac{\ell}{n_0} X_\ell\| < \delta) > 0$ and $\frac{m}{n_0} B < \delta$. Then, for $n \geq m$:

$$P(\overline{X}_n \in A) \geq P(\frac{mq_n}{n}\overline{X}_{mq_n} \in F^\delta)P(\|\frac{r_n}{n}\overline{X}_{r_n}\| < \delta) \quad ,$$

where $q_n = [n/m]$ and $r_n = n - [n/m]m$. Note that for $n \geq n_0$, $\overline{X}_{mq_n} \in F$ implies that $\mathrm{dist}(\frac{mq_n}{n}\overline{X}_{mq_n},F) \leq \frac{r_n}{n} B < \delta$; and so $P(\frac{mq_n}{n}\overline{X}_{mq_n} \in F^\delta) \geq P(\overline{X}_{mq_n} \in F) \geq P(\overline{X}_m \in F)^{q_n}$. Hence, for $n \geq n_0$:

$$P(\overline{X}_n \in A) \geq \alpha P(\overline{X}_m \in F)^{q_n} > 0 \quad . \quad \blacksquare$$

Given $n \geq 1$, define $\mu_n = P \circ (\overline{X}_n)^{-1}$ and set $\mu = \mu_1$. For $A \in \mathcal{B}_E$, set $\underline{\ell}(A) = \underline{\lim}_{n\to\infty} \frac{1}{n}\log \mu_n(A)$, $\overline{\ell}(A) = \overline{\lim}_{n\to\infty} \frac{1}{n}\log \mu_n(A)$, and $\ell(A) = \underline{\ell}(A)$ if $\underline{\ell}(A) = \overline{\ell}(A)$. The following relations are obvious:

$$\underline{\ell}(A) \leq \underline{\ell}(B) \text{ and } \overline{\ell}(A) \leq \overline{\ell}(B) \text{ if } A \subseteq B \quad . \tag{3.13}$$

We also have:

$$\underline{\ell}(A) \vee \underline{\ell}(B) \leq \underline{\ell}(A \cup B) \leq \overline{\ell}(A \cup B) \leq \overline{\ell}(A) \vee \overline{\ell}(B) \quad . \tag{3.14}$$

The left side of (3.14) is obvious from the first part of (3.13). To prove the right side, we assume that $\overline{\ell}(A) \vee \overline{\ell}(B) < \infty$. Given $\ell > \overline{\ell}(A) \vee \overline{\ell}(B)$, we have $\mu_n(A\cup B) \leq \mu_n(A) + \mu_n(B) \leq 2e^{n\ell}$ for large n's.

Hence, $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(A \cup B) \leq \ell$.

Next, note that, by Lemma 3.12, $\ell(A)$ exists for all open convex sets A . We now define

$$\lambda(x) = -\inf\{\ell(A) : A \text{ is open, convex and contains } x\} .$$

Clearly, $\lambda(x) = -\lim_{\delta\downarrow 0} \ell(B(x,\delta))$.

(3.15) <u>Lemma:</u> λ is lower semi-continuous and convex.

<u>Proof</u>: Let $c \in R^1$ be given. If $\lambda(x) > c$, then there exists an open convex $A \ni x$, and so $\ell(A) < -c$. Hence, $-\lambda(y) < -c$ for all $y \in A$, and so $A \subseteq \{\lambda > c\}$. This proves that λ is l.s.c.

To prove that λ is convex, note that (since λ is l.s.c.) it suffices to prove that $\lambda\left(\frac{x+y}{2}\right) \leq 1/2(\lambda(x) + \lambda(y))$. Now, let $x \neq y$ be given and let A be an open convex neighborhood of $\frac{x+y}{2}$. Choose open convex $B \ni x$ and $C \ni y$ so that $\frac{B+C}{2} \subseteq A$. Then, $\mu_n(B)\mu_n(C) \leq \mu_{2n}(A)$; and so $\frac{1}{2n} \log \mu_{2n}(A) \geq 1/2\left(\frac{1}{n} \log \mu_n(B) + \frac{1}{n} \log \mu_n(C)\right)$. Passing to the limit as $n \to \infty$, we arrive at $\ell(A) \geq 1/2(\ell(B) + \ell(C)) \geq -\frac{1}{2}(\lambda(x)+\lambda(y))$. Since this holds for all open convex $A \ni \frac{x+y}{2}$, we now get $-\lambda\left(\frac{x+y}{2}\right) \geq -1/2(\lambda(x)+\lambda(y))$. ∎

(3.16) <u>Theorem</u>: If A is an open set, then $\underline{\ell}(A) \geq -\inf_{x\in A} \lambda(x)$. If A is a compact set, then $\overline{\ell}(A) \leq -\inf_{x\in A} \lambda(x)$. Finally, if A is the finite union of open convex sets, then $\underline{\ell}(A) = -\inf_{x\in A} \lambda(x) = \overline{\ell}(A)$ (i.e. $\ell(A)$ exists and equals $-\inf_{x\in A} \lambda(x)$.)

<u>Proof</u>: Suppose A is open. Given $x \in A$, choose an open convex $B \ni x$ so that $B \subseteq A$. Then $-\lambda(x) \leq \ell(B) \leq \underline{\ell}(A)$. Thus, $\underline{\ell}(A) \geq -\inf_{x\in A} \lambda(x)$.

Next, suppose that A is compact. Given $c > -\inf_{x\in A} \lambda(x)$, choose, for each $x \in A$, an open convex neighborhood $B(x)$ so that $\ell(B(x)) < c$. Select a finite set $x_1, \ldots, x_N \in A$ so that $A \subseteq \bigcup_1^N B(x_k)$. Then:

$$\begin{aligned}\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(A) &\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log[N(\max_{1\leq k\leq N} \mu_n(B(x_k)))] \\ &\leq \max_{1\leq k\leq N} \overline{\lim} \frac{1}{n} \log(\mu_n(B(x_k))) \\ &< c \ .\end{aligned}$$

This proves that $\overline{\lim_{n\to\infty}} \frac{1}{n} \mu_n(A) \leq -\inf_{x\in A} \lambda(x)$ for compact A.

Finally, suppose that $A = \bigcup_1^N A_k$, where each A_k is open and convex. Then, by (3.14) and the fact that $\underline{\ell}(A_k) = \overline{\ell}(A_k)$ for $1 \leq k \leq N$, we see that $\ell(A)$ exists and equals $\max_{1\leq k\leq N} \ell(A_k)$. Thus, we will be done as soon as we show that $\ell(A) = -\inf_{x\in A} \lambda(x)$ for open convex A's. Hence, we assume that A itself is open and convex. Since, $\ell(A) \geq -\lambda(x)$, $x \in A$, we certainly have that $\ell(A) \geq -\inf_{x\in A} \lambda(x)$. Thus, we need only check that $\ell(A) \leq -\inf_{x\in A} \lambda(x)$; and, in doing so, we may and will assume that $\ell(A) > -\infty$. Given $\varepsilon > 0$, choose $N \geq 1$ so that $\frac{1}{n} \log \mu_n(A) \geq \ell(A) - \varepsilon$ whenever $n \geq N$. Then (since E is a Polish space) we can find a $K \subset\subset A$ so that $\frac{1}{N} \log \mu_N(A) - \frac{1}{N} \log \mu_N(K) < \varepsilon$. We want to show that K may be chosen to be convex as well as compact. To this end, cover K with a finite number L of open balls $B_1, \ldots, B_L$ such that $\overline{B}_\ell \subseteq A$, $1 \leq \ell \leq L$. Set $K_\ell = K \cap \overline{B}_\ell$. Then K_ℓ is compact and so, by Mazur's theorem, the closed convex hull $\hat{K}_\ell$

of K_ℓ is also compact. Clearly $\hat{K}_\ell \subseteq \overline{B}_\ell$. Thus $\Gamma \equiv \{\sum_1^L \alpha_\ell x_\ell : \sum_1^L \alpha_\ell = 1$; $\alpha_1 \wedge \cdots \wedge \alpha_L \geq 0$; and $x_\ell \in \hat{k}_\ell$, $1 \leq \ell \leq L\} \subseteq A$ and contains K . Moreover, Γ is closed (in fact, compact) and convex. Thus, the closed convex hull $\hat{K}$ of K is both compact and contained in A . In other words, we may and will assume that K itself is convex.

Now, set $f(n) = -\log \mu_n(K)$. Then f is subadditive and $f(N) \leq -(\ell(A)-2\varepsilon)N$. Hence, $f(mN)/mN \leq -\ell(A) + 2\varepsilon$ for all $m \geq 1$; and so $-\overline{\ell}(K) = \underline{\lim}_{n\to\infty} f(n)/n \leq -\ell(A) + 2\varepsilon$. Since $\overline{\ell}(K) \leq -\inf_{x\in K} \lambda(x)$, we have therefore proved that:

$$\ell(A) - 2\varepsilon \leq \overline{\ell}(K) \leq -\inf_{x\in K} \lambda(x) \leq -\inf_{x\in A} \lambda(x) \quad . \quad \square$$

(3.17) <u>Corollary</u>: For each $x \in E$:

$$\lambda(x) = -\inf\{\ell(H) : H \text{ is an open half-space and } H \ni x\} \quad .$$

<u>Proof</u>: Set $\tilde{\lambda}(x)$ equal to the right hand side of the above equation. Clearly $-\tilde{\lambda}(x) \geq -\lambda(x)$, and so $\tilde{\lambda}(x) \leq \lambda(x)$. To prove the opposite inequalty, let $c < \lambda(x)$ be given and set $C = \{y : \lambda(y) \leq c\}$. Then, by (3.15) , C is closed and convex. Clearly, $x \notin C$. Thus, by the Hahn-Banach theorem, there is an open half-space $H \ni x$ such that $H \cap C = \emptyset$. Note that for all $y \in H$, $\lambda(y) > c$. Thus, by the last part of Theorem (3.16) , $-\ell(H) = \inf_{y\in H} \lambda(y) \geq c$; and so, $\tilde{\lambda}(x) \geq -\ell(H) \geq c$. $\square$

The importance of (3.17) is that it allows us to reduce the calculations of $\lambda(x)$ to 1-dimensional computations. To see how this is done, let E^* be the dual space of E . Given $x^* \in E^*$, let μ^{x^*} be the distribution on R^1 of $x \to x^*(x)$ under μ . Define $\ell^{x^*}(\Lambda)$, for open convex $\Lambda \subseteq R^1$, and $\lambda^{x^*}(\eta)$, for $\eta \in R^1$, corresponding to μ^{x^*} .

(3.18) Lemma: For each $x \in E$, $\lambda(x) = \sup_{x^* \in E^*} \lambda^{x^*}(x^*(x))$.

Proof: Let $x^* \in E^*$ be given. By (3.17) applied to λ^{x^*}, we have:

$$-\lambda^{x^*}(x^*(x)) = \inf\{\ell^{x^*}((x^*(x)-\varepsilon,\infty)) \wedge \ell^{x^*}((-\infty,x^*(x)+\varepsilon)) : \varepsilon > 0\} .$$

For $\varepsilon > 0$, define $H_{\pm}(x^*,\varepsilon) = \{y \in E : \pm x^*(y) > \pm(x^*(x) \mp \varepsilon)\}$. Then $\ell^{x^*}((x^*(x)-\varepsilon,\infty)) = \lim_{n\to\infty} \frac{1}{n} \log \mu_n(H_+(x^*,\varepsilon)) = \ell(H_+(x^*,\varepsilon))$; and similarly, $\ell^{x^*}((-\infty,x^*(x)+\varepsilon)) = \ell(H_-^*(x^*,\varepsilon))$. Thus

$$-\lambda^{x^*}(x^*(x)) = \inf_{\varepsilon>0} \ell(H_+(x^*,\varepsilon)) \wedge \ell(H_-(x^*,\varepsilon)) .$$

But if H is any open half space containing x, then (by the Hahn-Banach theorem) $H = H_+(x^*,\varepsilon)$ for some $x^* \in E^*$ and $\varepsilon > 0$. Thus, by (3.17) applied to λ, $-\lambda(x) = \inf\{\ell(H_+(x^*,\varepsilon)) : x^* \in E$ and $\varepsilon > 0\}$. Since $H_-(x^*,\varepsilon) = H_+(-x^*,\varepsilon)$, this completes the proof. □

(3.19) Lemma: Let μ be a probability measure on R^1 satisfying $M(\xi) = \int e^{\xi y}\mu(dy) < \infty$ for all $\xi \in R^1$. Define λ and I for μ . Then $\lambda = I$.

Proof: We first show that $\lambda \leq I$. Indeed, $-\lambda(x) = \lim_{\varepsilon\downarrow 0} \ell((x-\varepsilon,x+\varepsilon)) \geq -\lim_{\varepsilon\downarrow 0} \inf_{|y-x|<\varepsilon} I(y) \geq -I(x)$, where we have used Theorem (3.8) to get $\ell((x-\varepsilon,x+\varepsilon)) \geq -\inf_{|y-x|<\varepsilon} I(y)$. To prove the opposite inequality, set $a = \int y\mu(dy)$. By the law of large numbers, $\ell((a-\varepsilon,a+\varepsilon)) = 0$ for all $\varepsilon > 0$. Thus $\lambda(a) = 0$. By Lemma (3.3) , $I(a) = 0$. Thus $\lambda(a) = I(a)$. Next, suppose that $x > a$. Then, for $0 < \varepsilon < x-a$:

$$\ell((x-\varepsilon,x+\varepsilon)) \le \overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n([x-\varepsilon,\infty)) \le -I(x-\varepsilon)$$

(since by (3.8), $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n([x-\varepsilon,\infty)) \le -\inf_{y\ge x-\varepsilon} I(y)$ and, by (3.3),

$I(x-\varepsilon) = \inf_{y\ge x-\varepsilon} I(y)$). Hence, $\lambda(x) \ge \overline{\lim_{\varepsilon\downarrow 0}} I(x-\varepsilon) \ge \underline{\lim_{y\to x}} I(y) \ge I(x)$, since I is l.s.c. A similar argument applies to $x \in (-\infty,a)$. □

(3.20) <u>Theorem</u>: Let μ be a probability measure on the separable Banach space E satisfying $M(x^*) = \int \exp(x^*(x))\mu(dx) < \infty$ for all $x^* \in E^*$. Then, for all open convex $A \subseteq E$, $\ell_\mu(A) = \lim_{n\to\infty} \frac{1}{n} \log \mu_n(A)$ exists. Moreover, if $\lambda_\mu(x) = -\inf\{\ell_\mu(A) : A \ni x$ such that A is open and convex$\}$, then λ_μ is a l.s.c. convex function and

(3.21) $$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(K) \le -\inf_{x\in F} \lambda_\mu(K) \quad , \quad K \text{ compact}$$

(3.22) $$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(G) \ge -\inf_{x\in F} \lambda_\mu(G) \quad , \quad G \text{ open,}$$

and

$$\lim_{n\to\infty} \frac{1}{n} \log \mu_n(A) = -\inf_{x\in A} \lambda_\mu(x) \quad , \quad A \text{ open and convex.}$$

Finally, if $I_\mu(x) = \sup_{x^*\in E^*} (x^*(x) - \log M(x^*))$, $x \in E$, then $I_\mu = \lambda_\mu$.

<u>Proof</u>: The only statement not proved already is the final one. But this result is an easy consequence of (3.18) and (3.19). □

(3.23) <u>Exercise</u>: Let $E = C([0,T];R^d)$ and let μ be Wiener measure on E (i.e. imbed E in $\Omega = C([0,\infty);R^d)$ in the obvious way and let μ

be the restriction of $\mathfrak{W}$ to $\mathcal{M}_T$). Show that $\log M_\mu(\gamma) = 1/2\int_0^T|\gamma(t)|^2 dt$, where $\gamma \in C([0,T];R^d)^*$ and $\gamma(t) = \gamma((t,T])$. Next, show that

$$I_\mu(\phi) = \begin{cases} \infty & \text{if } \phi(0) \neq 0 \text{ or } \phi \text{ is not absolutely continuous} \\ 1/2 \int_0^T |\phi(t)|^2 dt & \text{otherwise.} \end{cases}$$

Finally, use the results of this section to recover Schilder's theorem (i.e. Theorem (1.16)), at least for closed sets F in that theorem which are compact.

(3.24) Remark: Before proceeding, we point out that we need not restrict ourselves to separable Banach spaces E . Indeed, let E be a locally convex, Hausdorff topological vector space and suppose that H is a closed convex subset of E such that the induced topology on H admits a metric ρ having the properties that

i) (H,ρ) is a complete separable metric space,

ii) if $N \geq 2$, $\{x_n\}_1^N \cup \{y_n\}_1^N \subseteq H$, and $\{\alpha_n\}_1^N \subseteq (0,1)$ satisfy $\sum_1^N \alpha_i = 1$, then

$$\rho(\sum_1^N \alpha_n x_n, \sum_1^N \alpha_n y_n) \leq \max_{1\leq n\leq N} \rho(x_n,y_n) .$$

then the preceding result continues to hold when μ is a probability measure on $(E,\mathcal{B}_E)$ such that $\mu(H) = 1$. In order to carry out this generalization, one needs to handle two points. The first point is rather pedantic; namely, we must make sure that we can add our random variables. To handle this point, suppose $X_1,\ldots,X_n$ are E-valued random variables on the complete probability space $(\Omega,\mathcal{M},P)$, and assume that $P(X_m \notin H) = 0$, $1 \leq m \leq n$. Then, in order to prove that $\sum_1^n X_m$ is $\mathcal{M}$-measurable as an E-valued random variable, we may and will assume that $X_m(\omega) \in H$, $1 \leq m \leq n$ and $\omega \in \Omega$. Clearly, $\omega \to (X_1(\omega),\ldots,X_n(\omega)) \in H^n$ is $\mathcal{M}$-measurable into $(H^n,(\mathcal{B}_H)^n)$. Also,

$(x_1,\ldots,x_n) \in H^n \to \sum_1^n x_m \in E$ is continuous and therefore $\mathcal{B}_{H^n}$-measurable into E. Finally, since H is Polish, $\mathcal{B}_{H^n} = (\mathcal{B}_H)^n$. Combining these, we conclude that $\omega \to \sum_1^n X_m(\omega) \in E$ is $\mathcal{M}$-measurable.

The second point is a little more interesting. Namely, we must learn how to generalize the argument used at the end of the proof of Theorem (3.16). The only difficulty in doing so is overcome by the following lemma.

(3.25) <u>Lemma:</u> Let E and H be as in the preceding and suppose that K is a compact set contained in H. Then, the closed convex hull $\hat{K}$ of K is compact. In particular, if ν is a probability measure on E supported on H and if A is a convex open set, then for each $\varepsilon > 0$ there is a compact convex $K \subseteq A$ such that $\nu(A) - \nu(K) < \varepsilon$.

<u>Proof:</u> Suppose that we have proved the first part. To prove the second part from the first, we proceed as follows. Since $A \cap H$ is Polish, and $\nu(A) = \nu(A \cap H)$, given $\varepsilon > 0$ we can find a compact $K \subseteq A \cap H$ such that $\nu(A) - \nu(K) < \varepsilon$. Now, using the first part in place of Mazur's theorem, we can argue, in precisely the same way as we did in the proof of Theorem (3.16), that $\hat{K}$ not only is compact but also is contained in A.

To prove the first part, note that we will know that $\hat{K}$ is compact as soon as we show that it is ρ-totally bounded. But, for give $\delta > 0$, there exists $a_1,\ldots,a_N \in K$ such that $K \subseteq \bigcup_1^N B(a_N,\delta)$ $(B(a,r) \equiv \{x \in H\colon \rho(x,a) < r\})$. Clearly $\hat{K} \subseteq \Gamma \equiv \{\sum_1^N \alpha_n x_n : \{\alpha_n\}_1^N \subseteq [0,1]$, $\sum_1^N \alpha_n = 1$, and $x_n \in \overline{B(a_n,\delta)}$ for $1 \le n \le N\}$. At the same time, our assumptions about ρ tell us that $\Gamma \subseteq \{x \in H\colon \rho(x, \sum_1^N \alpha_n a_n) \le \delta$ for some $\{\alpha_n\}_1^N \subseteq [0,1]$ satisfying $\sum_1^N \alpha_n = 1\}$. Since $\{\sum_1^N \alpha_n a_n\colon \{\alpha_n\}_1^N \subseteq [0,1]$ and $\sum_1^N \alpha_n = 1\} \subset\subset H$, it is now clear that Γ, and therefore K, admits a finite covering by balls of radius $< 2\delta$. □

(3.24) Continued: Having handled these two points, there are no other obstacles preventing us from repeating the arguments leading to Theorem (3.20) for E and μ which satisfy the hypotheses in (3.25). The main reason for our interest in this generalization is the following. Let Γ be a Polish space and denote by $\mathcal{M}(\Gamma)$ the space of totally finite signed measures on Γ. Give $\mathcal{M}(\Gamma)$ the topology of weak convergence (i.e. $\nu_\alpha \to \nu$ if and only if $\int f d\nu_\alpha \to \int f d\nu$, $f \in C_b(\Gamma)$). Clearly this makes $\mathcal{M}(\Gamma)$ into a locally convex, Hausdorff topological vector space. Moreover, if $\mathcal{M}_1(\Gamma)$ denotes the space of probability measures on Γ, then $\mathcal{M}(\Gamma)$ is a closed convex subset of $\mathcal{M}(\Gamma)$. Finally, if ρ is the Lévy metric on $\mathcal{M}_1(\Gamma)$ (i.e., $\rho(\mu,\nu) \leq \varepsilon$ if and only if $\mu(F) \leq \nu(F^\varepsilon) + \varepsilon$ and $\nu(F) \leq \mu(F^\varepsilon) + \varepsilon$ for all closed F in Γ), then ρ satisfies the required conditions. Thus, we can take $E = \mathcal{M}(\Gamma)$ and $H = \mathcal{M}_1(\Gamma)$.

By far the most serious deficiencies in the line of reasoning which led us to Theorem (3.20) are that it fails to tell us when I_μ is a true rate function (i.e. $\{x : I_\mu(x) \leq L\}$ is compact, for all $L > 0$) and when $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} I_\mu(x)$ for all closed F (i.e. not just compact ones). The next result provides us with our first step toward removing these deficiencies.

(3.26) Theorem: Let E and μ be as above. Assume that for each $L > 0$ there is a compact $K_L \subset E$ such that $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(K_L^c) \leq -L$. Then for each $L > 0$, $\{x : \lambda_\mu(x) \leq L\}$ is a compact convex set. Moreover, for all closed F, $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} \lambda_\mu(F)$.

Proof: Since we already know that λ_μ is l.s.c. and convex, we will have proved that $\{x : \lambda_\mu(x) \leq L\}$ is compact and convex as soon as we show that it is contained in a compact set. But,

$-\inf_{x \notin K_{L+1}} \lambda_\mu(x) \leq \underline{\lim}_{n\to\infty} \frac{1}{n} \log \mu_n(K_{L+1}^c) \leq -L-1$. Thus $\lambda_\mu(x) \geq L+1$ for all $x \notin K_{L+1}$ and so $\{x : \lambda_\mu(x) \leq L\} \subseteq K_{L+1}$.

To prove the second part, set $F_L = F \cap K_L$. Then for all $\varepsilon > 0$ and $L > 0$, $\mu_n(F) \leq \mu_n(F_L) + \mu_n(K_L^c) \leq e^{-n((\ell-\varepsilon)\wedge 1/\varepsilon)} + e^{-nL/2} \leq$ $2\exp(-n((\ell-\varepsilon) \wedge 1/\varepsilon \wedge L/2))$ for sufficiently large n's , where $\ell \equiv \inf_{x\in F} \lambda_\mu(x) \leq \inf_{x\in F_L} \lambda_\mu(x)$. Hence

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -((\ell-\varepsilon) \wedge 1/\varepsilon \wedge L/2)$$

for all $L > 0$ and $\varepsilon > 0$. Clearly this leads to the desired estimate. □

(3.27) <u>Corollary</u>: Let E and μ be as (3.25) . Suppose that $\Phi : E \to [0,\infty) \cup \{\infty\}$ is a l.s.c., convex function satisfying $\{x : \Phi(x) \leq L\} \subset\subset E$ for all $L > 0$. If $\int e^{\varepsilon\Phi(x)}\mu(dx) < \infty$ for some $\varepsilon > 0$, then for all $L > 0$ there is a $K_L \subset\subset E$ such that $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(K_L^c) \leq -L$. In particular, $\{x : \lambda_\mu(x) \leq L\}$ is compact and convex for all $L > 0$ and $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(F) \leq -\inf_{x\in F} \lambda_\mu(x)$ for all closed F .

<u>Proof</u>: We need only prove the first part. Moreover, after replacing Φ by $\varepsilon\Phi$, we may and will assume that $\varepsilon = 1$.

Note that $\int e^{n\Phi(x)}\mu_n(dx) \leq (\int e^{\Phi(x)}\mu(dx))^n$, and so:

$$\mu_n(\{x : \Phi(x) > L\}) \leq e^{-nL}(\int e^{\Phi(x)}\mu(dx))^n \quad .$$

Hence, if $M = \log(\int e^{\Phi(x)}\mu(dx))$, then

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(\{x: \Phi(x) > L+M\}) \leq -L \ . \ \square$$

(3.28) Remark: When $E = R^d$ for some $d > 1$, the hypotheses of (3.27) are satisfied as soon as $\int e^{\varepsilon|x|}\mu(dx) < \infty$ for some $\varepsilon > 0$. In the case of Wiener measure (cf. (3.24)), we can take

$$\Phi(\psi) = |\psi(0)| + \sup_{0\le s<t\le T} \frac{|\psi(t)-\psi(s)|}{|t-s|^{1/4}}$$

(cf. for example, Lemma (8.7) in [S., San Flour].

We next want to show that the hypotheses of (3.27) are satisfied whenever E is a separable Banach space and μ satisfies $\int e^{\xi\|x\|_E}\mu(dx) < \infty$ for all $\xi > 0$.

(3.29) Lemma: Let E be a separable Banach space. Given a probability measure ν on E satisfying $\int\|x\|_E\nu(dx) < \infty$, there exists a unique $m(\nu) \in E$ such that $x^*(m(\nu)) = \int x^*(x)\nu(dx)$ for all $x^* \in E^*$. Moreover, if C is a closed convex set on which ν is supported, then $m(\nu) \in C$. Finally, if $B \subseteq \mathcal{M}_1(E)$ satisfies $\lim_{R\uparrow\infty}\sup_{\nu\in B}\int_{\|x\|_E>R}\|x\|_E\nu(dx) = 0$, then

$\lim_{R\uparrow\infty}\sup_{\nu\in\overline{B}}\int_{\|x\|_E>R}\|x\|_E\nu(dx) = 0$ and $\nu \to m(\nu)$ is continuous on $\overline{B}$.

Proof: Assume that $\int\|x\|_E\nu(dx) < \infty$. Then we can find a non-decreasing sequence of compacts $K_n \subseteq E$ such that $\int_{K_n^c}\|x\|_E\nu(dx) \le 1/n$. For each n, we can find a finite number N_n of disjoint measurable sets $A_{n,1},\ldots,A_{n,N_n} \subseteq K_n$ covering K_n such that for each $1 \le k \le N$ there is an $x_{n,k} \in E$ for which $A_{n,k} \subseteq B(x_{n,k},1/n)$. Define $y_n = \sum_{k=1}^{N_n} x_{n,k}\nu(A_{n,k})$. Then, for $m \le n$:

$$\|y_n - y_m\| \le \Big\|\sum_{\ell=1}^{N_n}\sum_{k=1}^{N_m}(x_{n,\ell} - x_{m,k})\nu(A_{n,\ell}\cap A_{m,k})\Big\|_E$$

$$+ \Big\|\sum_{\ell=1}^{N_n} x_{n,\ell}\nu(A_{n,\ell}\cap K_m^c)\Big\|_E$$

$$\leq 2/m + \sum_{\ell=1}^{N_n} \sup_{A_{n,\ell}} \|x\|_E \nu(A_{n,\ell} \cap K_m^c)$$

$$\leq 4/m + \int_{K_m^c} \|x\|_E \nu(dx) \leq 5/m \quad .$$

Hence, y_n converges in E to some $m(\nu) \in E$. Moreover,

$$\left|x^*(y_n) - \int x^*(x)\nu(dx)\right|$$

$$\leq \|x^*\|_{E^*} \sum_{\ell=1}^{N_n} \int_{A_{n,\ell}} \|x_{n,\ell} - x\| \nu(dx) + \|x^*\|_{E^*} \int_{K_n^c} \|x\|_E \nu(dx) \quad .$$

$$\to 0 \quad .$$

Thus, $x^*(m(\nu)) = \int x^*(x)\nu(dx)$, $x^* \in E^*$. The uniqueness of $m(\nu)$ is obvious. Moreover, if C is a closed convex set on which ν is supported and if $m(\nu) \notin C$, then, by the Hahn-Banach theorem, there is an $x^* \in E^*$ and a $c \in R^1$ such that $x^*(m(\nu)) > c \geq \sup_{x \in C} x^*(x)$. But then,

$$x^*(m(\nu)) > c \geq \int_C x^*(x)\nu(dx) = x^*(m(\nu)) \quad .$$

Hence, $m(\nu)$ must be in C .

Finally, suppose that $\lim_{R\uparrow\infty} \sup_{\nu\in B} \int_{\|x\|_E > R} \|x\|_E \nu(dx) = 0$. Since $\nu \to \int_{\|x\|_E>R} \|x\|_E \nu(dx)$ is l.s.c. for each $R > 0$, $\sup_{\nu\in\overline{B}} \int_{\|x\|_E>R} \|x\|_E \nu(dx) = \sup_{\nu\in B} \int_{\|x\|_E>R} \|x\|_E \nu(dx)$. Hence $\lim_{R\uparrow\infty} \sup_{\nu\in\overline{B}} \int_{\|x\|_E>R} \|x\|_E \nu(dx) = 0$. Now suppose that $\{\nu_n\}_1^\infty \subseteq \overline{B}$ and that $\nu_n \to \nu_0$. Given $\varepsilon > 0$, choose $R > 0$ so that $\sup_{\nu\in\overline{B}} \int_{\|x\|_E>R} \|x\|_E \nu(dx) < \varepsilon$. Choose $\rho \in C([0,\infty);[0,1])$ so that $\rho(\xi) = 1$, $\xi \leq R$, and $\rho(\xi) = 0$, $\xi \geq R+1$. For $x^* \in E^*$, define $F_{x^*}(x) = \rho(\|x\|)x^*(x)$. Then $\{F_{x^*} : \|x^*\|_{E^*} \leq 1\}$ is a uniformly bounded, equi-continuous family of functions on E into R^1 . Hence,

$$\lim_{n\to\infty} \sup_{\|x^*\|_{E^*}\le 1} \left|\int F_{x^*}(x)\nu_n(dx) - \int F_{x^*}(x)\nu_0(dx)\right| = 0 \quad ;$$

and so:

$$\overline{\lim_{n\to\infty}} \sup_{\|x^*\|_{E^*}\le 1} \left|\int x^*(x)\nu_n(dx) - \int x^*(x)\nu_0(dx)\right|$$

$$\le \overline{\lim_{n\to\infty}} \sup_{\|x^*\|_{E^*}\le 1} \left|\int F_{x^*}(x)\nu_n(dx) - \int F_{x^*}(x)\nu_0(dx)\right|$$

$$+ \overline{\lim_{n\to\infty}} \left|\int (1-\rho(\|x\|)\|x\|_E \nu_n(dx)\right|$$

$$+ \left|\int (1-\rho(\|x\|))\|x\|_E \nu_0(dx)\right| \le 2\varepsilon \quad .$$

Clearly, this proves that $m(\nu_n) \to m(\nu_0)$ in E . □

Given a l.s.c. $f : [0,\infty) \to [0,\infty) \cup \{\infty\}$ satisfying $\lim_{\xi\uparrow\infty} f(\xi)/\xi = \infty$ and an $L > 0$, set

$$\Gamma(f;L) = \{\nu \in \mathcal{M}_1(E) : \int f(\|x\|_E)\nu(dx) \le L\} \quad .$$

Since $\nu \to \int f(\|x\|)\nu(dx)$ is l.s.c., $\Gamma(f;L)$ is closed. Moreover, given $\varepsilon > 0$, choose R so that $f(\xi)/\xi \ge 1/\varepsilon$, $\xi \ge R$. Then,
$\sup_{\nu\in\Gamma(f;L)} \int_{\|x\|_E>R} \|x\|_E \nu(dx) \le \varepsilon \sup_{\nu\in\Gamma(f,L)} \int f(\|x\|_E)\nu(dx) \le \varepsilon L$. Hence
$\lim_{R\uparrow\infty} \sup_{\nu\in\Gamma(f;L)} \int_{\|x\|_E>R} \|x\|_E \nu(dx) = 0$. In particular, by Lemma (3.29) ,
$\nu \to m(\nu)$ is continuous on $\Gamma(f;L)$.

(3.30) <u>Lemma</u>: Let ν be a probability measure on R^1 and set $M_\nu(\xi) = \int e^{\xi x}\nu(dx)$, $\xi \in R^1$, and $I_\nu(x) = \sup_\xi(\xi x - \log M_\nu(\xi))$, $x \in R^1$.
Then:

<u>i)</u> if $M_\nu(\xi) < \infty$ for small $|\xi|$, then $I_\nu(x) \to \infty$ as $|x| \to \infty$;

<u>ii)</u> if $M_\nu(\xi) < \infty$ for all $\xi \in R^1$, then $I_\nu(x)/|x| \to \infty$ as $|x| \to \infty$;

<u>iii)</u> for any ν such that $\int |x| \nu(dx) < \infty$, $\int \exp(\varepsilon I_\nu(x))\nu(dx) < \infty$ for $0 \leq \varepsilon < 1$.

<u>Proof</u>: Suppose $M_\nu(\xi) < \infty$ for $|\xi| \leq \varepsilon$. Then $I_\nu(\pm x) \geq \varepsilon|x| - \log M_\nu(\pm\varepsilon) \to \infty$ as $|x| \to \infty$. Thus <u>i)</u> is proved. Next, suppose $M_\nu(\xi) < \infty$ for all $\xi \in R^1$. Given $L > 0$, $I_\nu(x) \geq Lx - \log M_\nu(L)$, $x \geq 0$, and so $\lim\limits_{x\to+\infty} I_\nu(x)/x \geq L$. Hence $\lim\limits_{x\to+\infty} I_\nu(x)/x = +\infty$. Similarly, $\lim\limits_{x\to-\infty} I_\nu(x)/-x = +\infty$. Thus <u>ii)</u> is proved.

To prove <u>iii)</u>, set $a = \int x\nu(dx)$ and let $x \geq a$ be given. Since $\nu([x,\infty)) \leq e^{-\xi x}\int e^{\xi y}\nu(dy)$ for all $\xi \geq 0$,

$$\nu([x,\infty)) \leq \exp(-\sup_{\xi\geq 0}(\xi x - \log M_\nu(\xi))) = e^{-I_\nu(x)}.$$

Next, note that (just as in the case when $M_\nu(\xi) < \infty$ for all $\xi \in R^1$) : I_ν is non-negative, l.s.c. and convex on R^1; $I_\nu(a) = 0$; and I_ν is non-decreasing on $[a,\infty)$. Set $b = \sup\{x \geq a : I_\nu(x) < \infty\}$. From the properties of I_ν on $[a,\infty)$, we see that I_ν is continuous on $[a,b)$, $I_\nu(b) = \lim\limits_{x\uparrow b} I_\nu(x)$ if $b > a$, and $I_\nu(x) = \infty$ on (b,∞). If $I_\nu(b) < \infty$, then, since $\nu((b,\infty)) = 0$, $\int_{[a,\infty)} e^{\varepsilon I_\nu(x)}\nu(dx) \leq e^{\varepsilon I_\nu(b)} < \infty$. If $I_\nu(b) = \infty$, then $b > a$ and $\nu([b,\infty)) = 0$, and so $\int_{[a,\infty)} e^{\varepsilon I_\nu(x)}\nu(dx) = \nu(\{a\}) + \int_{(a,b)} e^{\varepsilon I_\nu(x)}\nu(dx)$. Since I_ν is continuous and non-decreasing on (a,b), $\int_{(a,b)} e^{\varepsilon I_\nu(x)}\nu(dx) = \varepsilon\int_{(a,b)}\nu([x,b))e^{\varepsilon I_\nu(x)} I_\nu(dx) \leq \varepsilon\int_{(a,b)} e^{-(1-\varepsilon)I_\nu(x)} I_\nu(dx) \leq \frac{\varepsilon}{1-\varepsilon} < \infty$. Thus, in any case,

$\int_{[a,\infty)} e^{\varepsilon I_\nu(x)} \nu(dx) < \infty$. A similar line of reasoning leads to $\int_{(-\infty,a]} e^{\varepsilon I_\nu(x)} \nu(dx) < \infty$. □

(3.31) Lemma: Let μ be a probability measure on the separable Banach space E . Assume that $\int \exp(\xi\|x\|_E)\mu(dx) > \infty$, $\xi \in (0,\infty)$. Then there is a l.s.c. convex function $f : [0,\infty) \to [0,\infty)$ satisfying $\lim_{t\uparrow\infty} f(\xi)/\xi = \infty$ such that

$$\lim_{L\uparrow\infty} \sup_{n\geq 1} \frac{1}{n} \log \mu^n\Big(\frac{1}{n}\sum_1^n \delta_{x_m} \notin \Gamma(f;L)\Big) = -\infty \quad .$$

Proof: Let ν denote the distribution of $\|x\|_E$ under μ . Since $\int_{R^1} e^{\xi x}\nu(dx) < \infty$, $\xi \in R^1$, $I_\nu(y)/y \to \infty$ as $y \uparrow \infty$. Thus, if $a = \int \|x\|_E \mu(dx)$

and $f(y) = \begin{cases} 0 , & y \in [0,a] \\ I_\nu(y) \wedge y^2 , & y \in (a,\infty) \end{cases}$, then f is a l.s.c., convex function on $[0,\infty)$ into $[0,\infty)$ satisfying $\lim_{y\to\infty} f(y)/y = \infty$. Moreover, if ν_f is the distribution of $f(y)$ under ν , then $\int e^{\xi y}\nu_f(dy) < \infty$ for $|\xi| < 1$ and so $I_{\nu_f}(y) \to \infty$ as $y \to \infty$. Finally, if $L > \int y\nu_f(dy)$, then

$$\mu^n\Big(\frac{1}{n}\sum_1^n \delta_{x_m} \notin \Gamma(f;L)\Big) = (\nu_f)_n((L,\infty))$$

$$\leq e^{-nI_{\nu_f}(L)} .$$

(3.32) Lemma: Let μ be a probability measure on the Polish space Γ . Then for each $L > 0$ there is a compact set C_L in $\mathcal{M}_1(\Gamma)$ such that $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu^n\Big(\frac{1}{n}\sum_1^n \delta_{x_m} \notin C_L\Big) \leq -L$.

Proof: Let $\varepsilon \in (0,1/2]$ and $0 < \delta < \varepsilon$ satisfying $\varepsilon \log \varepsilon/\delta \geq \log 2$ be given. Choose $K(\delta) \subset\subset \Gamma$ so that $\mu(K(\delta)^c) < \delta$ and set $G(\varepsilon,\delta) = \{\nu \in \mathcal{M}_1(\Gamma) : \nu(K(\delta)^c) \leq \varepsilon\}$. Then

$$\mu^n\Big(\frac{1}{n}\sum_1^n \delta_{x_m} \notin G(\varepsilon,\delta)\Big) \leq (\delta/\varepsilon)^{n\varepsilon/2} . \tag{3.33}$$

To see (3.33) , note that:

$$\mu^n\Big(\frac{1}{n}\sum_1^n \delta_{x_m} \notin G(\varepsilon,\delta)\Big)$$

$$= \mu^n\Big(\frac{1}{n}\sum_1^n \chi_{K(\delta)^c}(x_m) > \varepsilon\Big)$$

$$= (\beta_\delta)_n((\varepsilon,\infty))$$

$$\leq e^{-nI_{\beta_\delta}(\varepsilon)}$$

where $\beta_\delta = p_\delta \delta_{\{1\}} + (1-p_\delta)\delta_{\{0\}}$ and $p_\delta = \mu(K(\delta)^c) < \delta$. Using (3.10) i), we have

$$I_{\beta_\delta}(\varepsilon) = \varepsilon \log \varepsilon/p_\delta + (1-\varepsilon)\log \frac{1-\varepsilon}{1-p_\delta}$$

$$\geq \varepsilon \log \varepsilon/\delta - (1-\varepsilon)\log \frac{1}{1-\varepsilon} \geq \varepsilon \log \varepsilon/\delta - 1/2 \log 2$$

$$\geq \varepsilon/2 \log \varepsilon/\delta \quad ,$$

since $x \to (1-x) \log \frac{1}{1-x}$ is non-decreasing on $[0,1/2]$. Clearly (3.33) follows from this.

We now proceed as follows. Given $L > \log 2$ and $\varepsilon_\ell \downarrow 0$ with $\varepsilon_1 = 1/2$, set $\delta_\ell = \varepsilon_\ell e^{-2\ell L/\varepsilon_\ell}$ and choose $K_\ell \subset\subset \Gamma$ so that $\mu(K_\ell^c) < \delta_\ell$. Set $G_\ell = \{\nu \in \mathcal{M}_1(\Gamma) : \nu(K_\ell^c) \leq \varepsilon_\ell\}$. Then $C_L = \bigcap_{\ell=1}^{\infty} G_\ell$ is compact in $\mathcal{M}_1(\Gamma)$. Moreover, $(\varepsilon_\ell/\delta_\ell)^{\varepsilon_\ell} = e^{2\ell L} \geq 2$, and so, by (3.33) , $\mu^n(\frac{1}{n}\sum_1^n \delta_{x_m} \notin G_\ell) \leq (\delta_\ell/\varepsilon_\ell)^{n\varepsilon_\ell/2} \leq (e^{-nL})^\ell$. Hence, $\mu^n(\frac{1}{n}\sum_1^n \delta_{x_m} \notin C_L) \leq \sum_{\ell=1}^{\infty} (e^{-nL})^\ell \leq 2e^{-nL}$. $\square$

(3.34) <u>Theorem:</u> (Donsker & Varadhan): Let μ be a probability measure on the separable Banach space E . Assume that $\int e^{\xi\|x\|_E}\mu(dx) < \infty$ for all $\xi > 0$. Then, for each $L > 0$ there is a $K_L \subset\subset E$ such that $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(K_L^c) \leq -L$. In particular, I_μ is a convex rate function and $\{\mu_n : n \geq 1\}$ satisfies the large deviation principle with respect to I_μ .

Proof: Choose f as in Lemma (3.31). Then, by Lemma (3.31) and (3.32), for each $L > 0$ we can find an $R_L > 0$ and a $C_L \subset\subset \mathcal{M}_1(\Gamma)$ such that $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu^n(\frac{1}{n}\sum^n \delta_{x_m} \notin \Gamma(f;R_L) \cap C_L) \leq -L$. Now set $K_L = \{m(\nu) : \nu \in \Gamma(f;R_L) \cap C_L\}$. Since $\Gamma(f;R_L) \cap C_L \subset\subset \mathcal{M}_1(\Gamma)$ and $\nu \to m(\nu)$ is continuous on $\Gamma(f;R_L)$, $K_L \subset\subset E$. Finally, $\frac{1}{n}\sum_1^n x_m = m(\frac{1}{n}\sum_1^n \delta_{x_m})$ and therefore: $\mu_n(K_L^c) = \mu^n(m(\frac{1}{n}\sum_1^n \delta_{x_m}) \notin K_L) \leq \mu^n(\frac{1}{n}\sum_1^n \delta_{x_m} \notin \Gamma(f,R_L) \cap C_L)$. Thus $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu_n(K_L^c) \leq -L$. The rest of the theorem now follows from Theorems (3.20) and (3.26). □

Having devoted so much effort to derive a large deviations result associated with the law of large numbers for Banach space valued random variables, it seems only right to see that we have in fact proved the strong law of large numbers for such random variables.

(3.35) Theorem (Ranga Rao): Let $(\Omega,\mathcal{M},P)$ be a probability space and let E be a separable Banach space. Suppose that $X_1,\ldots,X_n,\ldots$ are independent, identically distributed E-valued random variables on $(\Omega,\mathcal{M},P)$. If $E^P[\|X_1\|_E] < \infty$, then $\overline{X}_n \xrightarrow{E} m(\mu)$ (a.s., P).

Proof: First suppose that $\|X_1\| \leq M$ (a.s., P) for some $M < \infty$. Let μ be the distribution of X_1. Then $\int e^{\xi\|x\|_E}\mu(dx) \leq e^{\xi M} < \infty$, $\xi > 0$; and so, by Theorem (3.34), there is a $K \subset\subset E$ such that $P(\overline{X}_n \notin K) = \mu_n(K^c) \leq Ce^{-n}$, $n \geq 1$, for some $C < \infty$. Hence $P((\exists m \geq 1)(\forall n \geq m)\overline{X}_n \notin K) = 0$; and therefore there is a $\Lambda_1 \in \mathcal{M}$ such that $P(\Lambda_1^c) = 0$ and, for each $\omega \in \Lambda_1$, $\{\overline{X}_n(\omega) : n \geq 1\}$ is relatively compact. Next note that, by the strong law for

R^1-valued random variables, for each $x^* \in E^*$: $x^*(\bar{X}_n) \to x^*(m(\mu))$ (a.s., P) . Since E is separable, and therefore E^* is separable in the weak* topology, we see that there is a $\Lambda_2 \in \mathcal{M}$ such that $P(\Lambda_2^c) = 0$ and $\bar{X}_n(\omega) \to m(\mu)$ weakly for all $\omega \in \Lambda_2$. But this means that $\bar{X}_n(\omega) \to m(\mu)$ strongly for all $\omega \in \Lambda_1 \cap \Lambda_2$.

We have now proved the theorem when $\|X_1\|_E$ is P-almost surely bounded. To handle the general case, set $X_n^{(M)} = \chi_{[0,M)}(\|X_n\|)X_n$, and $Y_n^{(M)} = X_n - X_n^{(M)}$ for $n \geq 1$ and $M \in (0,\infty)$. Given $\varepsilon > 0$, choose $M \in (0,\infty)$ so that $E^P[\|Y_1^{(M)}\|_E] < \varepsilon/4$. Next, choose N so that $P(\sup_{n\geq N} \frac{1}{n} \sum_{m=1}^{n} \|Y_m^{(M)}\|_E \geq \varepsilon/3) < \varepsilon/2$ (this can be done by the strong law for $\{\|Y_n^{(M)}\|_E : m \geq 1\}$) and $P(\sup_{n\geq N}\|\bar{X}_n^{(M)} - m(\mu^{(M)})\|_E \geq \varepsilon/3) < \varepsilon/2$, where $\mu^{(M)}$ is the distribution of $X_1^{(M)}$. Noting that $\|m(\mu) - m(\mu^{(M)})\|_E \leq E^P[\|Y_1^{(M)}\|]$ $< \varepsilon/3$, we now have:

$$P(\sup_{n\geq N}\|\bar{X}_n - m(\mu)\|_E \geq \varepsilon)$$

$$\leq P(\sup_{n\geq N}\|\bar{X}_n - \bar{X}_n^{(M)}\|_E \geq \varepsilon/3) + P(\sup_{n\geq N}\|\bar{X}_n^{(M)} - m(\mu^{(M)})\|_E \geq \varepsilon/3)$$

$$< P(\sup_{n\geq N} \frac{1}{n} \sum_{1}^{n} \|Y_m^{(M)}\|_E \geq \varepsilon/3) + \varepsilon/2$$

$$= \varepsilon \quad .$$

(3.36) <u>Exercise</u>: The proof of the strong law really requires much less than we have used. Indeed, show, directly from from the real-valued case, that $P(\frac{1}{n} \sum_{1}^{n} \delta_{X_n} \to \mu$ weakly$) = 1$. Next show that if $E^P[\|X_1\|_E] < \infty$, then there is a l.s.c. $f : [0,\infty) \to [0,\infty)$ such that

$\lim_{L\uparrow\infty} P((\exists m \geq 1)(\forall n \geq m)\frac{1}{n}\sum_1^n \delta_{X_\ell} \notin \Gamma(f;L)) = 0$. Combining these, show that $P(\{\overline{X}_n\}_1^\infty$ is relatively compact in E) = 1 and conclude that $\overline{X}_n \to m(\mu)$ (a.s., P) .

Related to the preceding considerations is a theorem, which in the one dimensional case, is due to Sanov.

(3.37) Lemma: Let Γ be a Polish space and $\alpha \in \mathcal{M}_1(\Gamma)$. Define $I_\alpha(\nu)$ for $\nu \in \mathcal{M}(\Gamma)$ by

$$I_\alpha(\nu) = \sup_{f\in C_b(\Gamma)} \left(\int f d\nu - \log\int e^f d\alpha\right) .$$

Then I_α is a convex rate function on $\mathcal{M}(\Gamma)$. Moreover,

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \alpha^n\left(\frac{1}{n}\sum_1^n \delta_{x_\ell} \in F\right) \leq -\inf_{\nu\in F} I_\alpha(\nu)$$

for closed sets F in $\mathcal{M}(\Gamma)$ and

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log \alpha^n\left(\frac{1}{n}\sum_1^n \delta_{x_\ell} \in G\right) \geq -\inf_{\nu\in G} I_\alpha(\nu)$$

for open sets G in $\mathcal{M}(\Gamma)$.

Proof: Set $E = \mathcal{M}(\Gamma)$, give E the weak topology, and denote by μ the distribution of δ_x under α . Since $\mathrm{supp}(\mu) \subseteq \mathcal{M}_1(\Gamma)$, we can use (3.24) to conclude that Theorem (3.20) applies to μ on E . Moreover, by Lemma (3.32) , the hypotheses of Theorem (3.26) are satisfied by μ on E . Finally, noting that $E^* = C_b(\Gamma)$ (indeed, if $\Lambda \in E^*$, set $f(x) = \Lambda(\delta_x)$, $x \in \Gamma$), we see from Theorem (3.20) that

$$I_\mu(\nu) = \sup_{f\in C_b(\Gamma)} \left(\int f d\nu - \log\int e^{\int f d\beta}\mu(d\beta)\right) = \sup_{f\in C_b(\Gamma)} \left(\int f d\nu - \log\int e^f d\alpha\right) .$$

(3.38) Lemma: Let everything be as in (3.37) . Then $I_\alpha(\nu) = \infty$ if

$\nu \notin \mathcal{M}_1(\Gamma)$ or $\nu \in \mathcal{M}_1(\Gamma)$ but $\nu \not\ll \alpha$, and $I_\alpha(\nu) = \int \phi \log \phi d\alpha$ if $\nu \in \mathcal{M}_1(\Gamma)$ and $d\nu = \phi d\alpha$.

<u>Proof</u>: First suppose that $\nu \notin \mathcal{M}_1(\Gamma)$. Then there is an open set $U \ni \nu$ such that $U \cap \mathcal{M}_1(\Gamma) = \emptyset$. Hence

$$I_\alpha(\nu) \geq \inf_{\beta \in U} I_\alpha(\beta) \geq -\underline{\lim}_{n\to\infty} \frac{1}{n} \log \alpha^n\Big(\frac{1}{n}\sum_1^n \delta_{x_\ell} \in U\Big)$$

$$= \infty \quad ,$$

since $\alpha^n\big(\frac{1}{n}\sum_1^n \delta_{x_\ell} \in U\big) \leq \alpha^n\big(\frac{1}{n}\sum_1^n \delta_{x_\ell} \notin \mathcal{M}_1(\Gamma)\big) = 0$.

Now suppose that $\nu \in \mathcal{M}_1(\Gamma)$ and that $d\nu = \phi d\alpha$. Since $\nu(\phi > 0) = 1$, we can choose $\varepsilon_0 > 0$ so that $\nu(\phi > \varepsilon_0) \geq 1/2$. For $0 < \varepsilon < \varepsilon_0$, set $d\nu_\varepsilon = (\chi_{(\varepsilon,\infty)} \circ \phi)d\nu/\nu(\phi > \varepsilon)$. Then, by Jensen's inequality, for each $f \in C_b(\Gamma)$ and $0 < \varepsilon < \varepsilon_0$:

$$\exp\Big(\int (f - \log\phi) d\nu_\varepsilon\Big) \leq \int e^f \frac{1}{\phi} d\nu_\varepsilon$$

$$= \frac{1}{\nu(\phi > \varepsilon)} \int_{\{\phi>\varepsilon\}} e^f d\alpha \quad .$$

Hence:

$$\frac{1}{\nu(\phi > \varepsilon)} \int_{\{\phi>\varepsilon\}} f d\nu - \log\Big(\int_{\{\phi>\varepsilon\}} e^f d\alpha\Big)$$

$$\leq \frac{1}{\nu(\phi > \varepsilon)} \int_{\{\phi>\varepsilon\}} \log\phi d\nu - \log\nu(\phi > \varepsilon) \quad .$$

Letting $\varepsilon \downarrow 0$, we get

$$\int f d\nu - \log\Big(\int e^f d\alpha\Big) \leq \int \log\phi d\nu$$

$$= \int \phi\log\phi d\alpha \quad .$$

Thus, $\nu \in \mathcal{M}_1(\Gamma)$ and $d\nu = \phi d\alpha$ imply that $I_\alpha(\nu) \leq \int \phi \log\phi d\alpha$.

It remains only to show that if $I_\alpha(\nu) < \infty$, then $\nu \in \mathcal{M}_1(\Gamma)$, $d\nu = \phi d\alpha$, and $I_\alpha(\nu) \geq \int \phi \log\phi d\alpha$. We already know that $\nu \in \mathcal{M}_1(\Gamma)$. Moreover, for all $f \in C_b(\Gamma)$:

$$\int f d\nu - \log\left(\int e^f d\alpha\right) \leq I_\alpha(\nu) \quad . \tag{3.39}$$

By tightness and Lusin's theorem, for bounded measurable ϕ and any $\varepsilon > 0$, there is a $K_\varepsilon \subset\subset \Gamma$ and an $f_\varepsilon \in C_b(\Gamma)$ such that $\nu(K_\varepsilon^c) \vee \alpha(K_\varepsilon^c) < \varepsilon$ and $f_\varepsilon\big|_{K_\varepsilon} = \phi\big|_{K_\varepsilon}$. Hence (3.39) continues to hold for all bounded measurable f. In particular, if $A \in \mathcal{B}_\Gamma$ has α measure 0, then (3.39) with $f = n\chi_A$ says that $n\nu(A) \leq I_\alpha(\nu)$. Hence $\nu << \alpha$. Set $\phi = \frac{d\nu}{d\alpha}$. If $\log\phi$ is bounded, then we can take $f = \log\phi$ in (3.39) and thereby obtain $\int \phi\log\phi d\alpha = \int \log\phi d\nu - \log\left(\int \phi d\alpha\right) \leq I_\alpha(\nu)$. To complete the proof, we must handle the case when $\log\phi$ is not bounded. To this end, first assume that $\phi \geq \theta > 0$. Then, with $f_n = \log(\phi \wedge n)$ in (3.39), we have

$$\int \phi\log\phi d\alpha = \int \log\phi d\nu = \lim_{n\to\infty} \int f_n d\nu$$

$$\leq \lim_{n\to\infty} \log\int e^{f_n} d\alpha + I_\alpha(\nu) = I_\alpha(\nu) \quad .$$

Finally, for $\theta \in [0,1]$, set $\nu_\theta = \theta\alpha + (1-\theta)\nu$. Then, $d\nu_\theta = \phi_\theta d\alpha$, where $\phi_\theta = \theta + (1-\theta)\phi \geq \theta$; and so, for $\theta \in (0,1)$, $\int \phi_\theta \log\phi_\theta d\alpha \leq I_\alpha(\theta\alpha + (1-\theta)\nu)$. Since $\theta \to I_\alpha(\theta\alpha + (1-\theta)\nu)$ is a l.s.c., bounded, convex function on $[0,1]$, $I_\alpha(\nu) = \lim_{\theta\downarrow 0} I_\alpha(\theta\alpha + (1-\theta)\nu) \geq \lim_{\theta\downarrow 0} \int \phi_\theta \log\phi_\theta d\alpha$. But

$$\int \phi_\theta \log\phi_\theta d\alpha = \theta\int \log\phi_\theta d\alpha + (1-\theta)\int \phi\log\phi_\theta d\alpha \quad ;$$

$$\theta\int \log\phi_\theta d\alpha \geq \theta\log\theta \to 0 \quad ;$$

and, since $t \to \log t$ is concave,

$$(1-\theta)\int\phi\log\phi_\theta d\alpha \geq (1-\theta)^2\int\phi\log\phi d\alpha \to \int\phi\log\phi d\alpha \quad .$$

Thus, $\int\phi\log\phi d\alpha \leq I_\alpha(\nu)$. □

Lemmas (3.37) and (3.38) combine to yield Sanov's theorem.

(3.40) <u>Theorem</u>: Let Γ be a Polish space and α a probability measure on Γ . Define $\Lambda_\alpha(\nu)$, $\nu \in \mathcal{M}(\Gamma)$, so that $\Lambda_\alpha(\nu) = \int(\log\frac{d\nu}{d\alpha})d\nu$ if $\nu \in \mathcal{M}_1(\Gamma)$ satisfies $\nu << \alpha$ and $\Lambda_\alpha(\nu) = \infty$ otherwise. Then Λ_α is a convex rate function on $\mathcal{M}(\Gamma)$. Moreover, if μ denotes the distribution of $x \to \delta_x$ under α on $\mathcal{M}(\Gamma)$, then $\{\mu_n : n \geq 1\}$ satisfies the large deviation principle with rate function Λ_α .

We finish this section with a discussion of Theorem 3.34 applied to the special case when μ is a <u>centered Gaussian</u>; that is, when for every element x^* of E^* the distribution of $x^*(x)$ under μ is Gaussian with mean 0 . In order to see that (3.14) applies to such μ's , we need the following result due to Skorohod.

(3.41) <u>Theorem</u>: Let μ be a centered Gaussian on the separable Banach space E . Then $\int e^{\xi\|x\|_E}\mu(dx) < \infty$ for all $\xi \in R^1$.

<u>Proof</u>: First, suppose that $\int e^{\varepsilon\|x\|_E}\mu(dx) < \infty$ for some $\varepsilon > 0$. Given $n \geq 1$, note that the distribution of $\frac{1}{n^{1/2}}\sum_1^n x_\ell$ under μ^n coincides with the distribution of x under μ . Thus:

$$\int e^{n^{1/2}\varepsilon\|x\|_E}\mu(dx) = \int\exp(n^{1/2}\frac{\varepsilon}{n^{1/2}}\|\sum_1^n x_\ell\|_E\,\mu^n(dx_1\ldots dx_n)$$

$$\leq (\int e^{\varepsilon\|x\|_E}\mu(dx))^n < \infty \quad .$$

In other words, it suffices to prove that $\int e^{\varepsilon\|x\|_E}\mu(dx) < \infty$ for some $\varepsilon > 0$.

In order to complete the proof, we first need an elementary fact about Brownian motion. Namely, given a d-dimension Brownian motion $(\beta(t),\mathcal{F}_t,P)$ on $(\Omega,\mathcal{F},P)$ starting at 0 and a continuous semi-norm $\|\cdot\|$ on R^d :

$$P(\sup_{0\le t\le T}\|\beta(t)\| \ge R) \le 2P(\|\beta(T)\| \ge R) \tag{3.42}$$

and

$$P(\sup_{0\le t\le T}\|\beta(t)\| \ge nR) \le P(\sup_{0\le t\le T}\|\beta(t)\| \ge R)^n \ , \ n \ge 1 \ , \tag{3.43}$$

for all $T > 0$ and $R > 0$. To prove (3.42) , let $\tau = \inf\{t \ge 0 : \|\beta(t)\| \ge R\}$. Set $C_R = \{\eta \in R^d : \|\eta\| \le R\}$ and make a measurable selection of $\eta \in \partial C_R \to \ell(\eta) \in R^d$ so that $|\ell(\eta)| = 1$ and $C_R \subseteq \{\xi \in R^d : (\ell(\eta),\xi-\eta)_{R^d} \le 0\}$ for all $\eta \in \partial C_R$. Then:

$$P(\|\beta(T)\| > R) \ge P((\beta(T)-\beta(\tau),\ell(\beta(\tau)))_{R^d} > 0 \ , \ \tau < T) \ .$$

By the strong Markov property (cf. Theorem (1.24))
$P((\beta(T)-\beta(\tau),\ell(\beta(\tau)))_{R^d} > 0 \mid \mathcal{F}_\tau) = P((\beta(T-\tau),\ell(\beta(\tau)))_{R^d} > 0)$ on $\{\tau < T\}$.
But $P((\beta(t),\ell)_{R^d} > 0) = 1/2$ for any $t > 0$ and $\ell \in R^d\backslash\{0\}$. Thus,

$$P(\|\beta(T)\| > R) \ge 1/2 \ P(\tau < T)$$

Since $P(\|\beta(T)\| = R) \ge P(\tau = T)$, $P(\|\beta(T)\| \ge R) \ge 1/2 \ P(\tau \le T)$; and this proves (3.42) . To prove (3.43) , set $\tau_n = \inf\{t \ge 0 : \|\beta(t)\| \ge nR\}$, $n \ge 1$. Then

$$P(\sup_{0\le t\le T}\|\beta(t)\| \ge (n+1)R)$$

$$\leq P(\sup_{\tau_n \leq t \leq T} \|\beta(t)-\beta(\tau_n)\| \geq R\ ,\ \tau_n < T) \quad .$$

By the strong Markov property, on $\{\tau_n < \infty\}$:

$$P(\sup_{\tau_n \leq t \leq T} \|\beta(t)-\beta(\tau_n)\| \geq R|\mathcal{F}_{\tau_n})$$

$$= P(\sup_{0 \leq t \leq T-\tau_n} \|\beta(t)\| \geq R)$$

$$\leq P(\sup_{0 \leq t \leq T} \|\beta(t)\| \geq R) \quad .$$

Thus

$$P(\sup_{0 \leq t \leq T} \|\beta(t)\| \geq (n+1)R)$$

$$\leq P(\sup_{0 \leq t \leq T} \|\beta(t)\| \geq R)P(\tau_n < T)$$

$$\leq P(\sup_{0 \leq t \leq T} \|\beta(t)\| \geq R)P(\sup_{0 \leq t \leq T} \|\beta(t)\| \geq nR) \quad .$$

Clearly (3.43) follows from this.

We can now proceed as follows. Given any finite set $\{x_1^*,\ldots,x_d^*\}$ contained in the closed unit ball of E^* , let A denote the covariance of $(x_1^*(x),\ldots,x_d^*(x))$ under μ (i.e. $A = ((E^{\mu}[x_i^*(x)x_j^*(x)]))_{1\leq i,j\leq d}$) and set $\sum = A^{1/2}$. If $(\beta(t),\mathcal{F}_t,P)$ is a d-dimensional Brownian motion starting at 0 , then the distribution of $(x_1^*(x),\ldots,x_d^*(x))$ under μ coincides with the distribution of $\sum(\beta(1))$ under P . Hence if $\|\eta\|_{\Sigma} = \max_{1\leq i\leq d}\left|\sum(\eta)_i\right|$, $\eta \in R^d$, then for any $R > 0$ and $n \geq 1$:

$$\mu(\{x : \max_{1\leq i\leq d}\left|x_i^*(x)\right| \geq nR\})$$

$$\leq P(\sup_{0 \leq t \leq 1} \|\beta(t)\|_{\Sigma} \geq nR)$$

$$\leq P(\sup_{0\leq t\leq 1} \|\beta(t)\|_\Sigma \geq R)^n$$

$$\leq (2P(\|\beta(1)\|_\Sigma \geq R))^n$$

$$= (2\mu(\{x : \sup_{1\leq i\leq d} |x_i^*(x)| \geq R))^n$$

$$\leq (2\mu(\{x: \|x\|_E \geq R\}))^n$$

In other words, for any finite subset F of the closed unit ball in E^*,

$$\mu(\{x : \max_{x^*\in F} |x^*(x)| \geq nR)$$

$$\leq (2\mu(\{x : \|x\|_E \geq R\}))^n$$

for all $n \geq 1$ and $R > 0$. But E is separable and so we can conclude that $\mu(\{x : \|x\|_E \geq nR\})\leq 2^n(\mu(\{x : \|x\|_E \geq R\}))^n$ for all $n \geq 1$ and $R > 0$. Since $\mu(\{x : \|x\|_E \geq R\}) \downarrow 0$ as $R \uparrow \infty$, it follows immediately that $E^\mu[e^{\varepsilon\|x\|_E}] < \infty$ for some $\varepsilon > 0$. □

As a consequence of Theorem (3.41) and (3.34), we see that if μ is a centered Gaussian measure on a separable Banach space E, then I_μ is a rate function and $\{\mu_n : n \geq 1\}$ satisfies the large deviation principle with respect to I_μ. We now want to find a more tractible expression for I_μ when μ is a centered Gaussian. In order to do so, we need to discuss some elementary facts about the structure of Gaussian measures on a Banach space.

Let E be a separable Banach space and $\rho : E^* \times E^* \to R^1$ a continuous, bilinear, symmetric, non-negative map. The triple (H,S,E) is said to be

<u>ρ-related</u> if H is a Hilbert space and $S : H \to E$ is a continuous, linear injection such that $\|S^*(x^*)\|_H^2 = \rho(x^*,x^*)$.

(3.44) <u>Lemma</u>: Suppose that (H,S,E) is ρ-related. Then $\sup_{x^* \in E^*} (x^*(x) - 1/2\, \rho(x^*,x^*)) < \infty$ implies that $x \in S(H)$. Moreover, if $h \in H$, then $\sup_{x^* \in E^*} (x^*(S(h)) - 1/2\, \rho(x^*,x^*)) = 1/2\|h\|_H^2$.

<u>Proof</u>: We first note that $S^*(E^*)$ is dense in H . Indeed, if $h \perp \overline{S^*(E^*)}$, then $x^*(S(h)) = (h,S^*(x^*))_H = 0$ for all $x^* \in E^*$; and so $S(h) = 0$. Since S is injective, this means that $h = 0$.

Next, suppose that $C = \sup_{x^* \in E^*} (x^*(x) - 1/2\, \rho(x^*,x^*)) < \infty$. Since $\rho(x^*,x^*) = \|S^*(x^*)\|_H^2$, this means that $x^*(x) - 1/2\|S^*(x^*)\|_H^2 \le C$ for all $x^* \in E^*$. In particular, if $S^*(x^*) = 0$, then $|\lambda x^*(x)| \le C$ for all $\lambda \in R^1$, and so $x^*(x) = 0$. This means that we can define $\Lambda : S^*(E^*) \to R^1$ by $\Lambda(S^*(x^*)) = x^*(x)$. Clearly Λ is linear. Moreover, $|\Lambda(S^*(x^*))| \le C + 1/2\|S^*(x^*)\|_H^2 \le C + 1/2$ for x^* satisfying $\|S^*(x^*)\|_H \le 1$. Thus, Λ is bounded. Since $S^*(E^*)$ is dense in H , there is a unique $h \in H$ such that $(h,S^*(x^*))_H = x^*(x)$, $x^* \in E^*$. In other words, $x = S(h)$.

To complete the proof, note that

$$\sup_{x^* \in E^*} (x^*(S(h)) - 1/2\, \rho(x^*,x^*))$$

$$= \sup_{x^* \in E^*} ((h,S^*(x^*))_H - 1/2\|S^*(x^*)\|_H^2)$$

$$= \sup_{g \in H}((h,g)_H - 1/2\|g\|_H^2) ,$$

where we have used the density of $S^*(E^*)$ in H to get the last equality. Finally:

$$(h,g)_H - 1/2\|g\|_H^2 \le \|h\|_H\|g\|_H - 1/2\|g\|_H^2$$
$$= -1/2(\|g\|_H - \|h\|_H)^2 + 1/2\|h\|_H^2$$
$$\le 1/2\|h\|_H^2 ,$$

and equality holds if $g = h$. □

(3.45) <u>Theorem</u>: Let E be a separable Banach space and let μ be a centered Gaussian measure on E . Define $\rho_\mu(x^*,y^*) = \int x^*(x)y^*(x)\mu(dx)$ for $(x^*,y^*) \in E^* \times E^*$. Then ρ_μ is a continuous, bilinear, symmetric, non-negative map. Next, let H_μ be the closure in $L^2(\mu)$ of the subspace $\{x^*(\cdot) : x^* \in E^*\}$; and, for $h \in H_\mu$, define $S_\mu(h) = \int xh(x)\mu(dx)$. (Note that $\int\|x\|_E|h(x)|\mu(dx) \le (\int\|x\|_E^2\mu(dx))^{1/2}(\int|h(x)|^2\mu(dx))^{1/2} < \infty$ and therefore $\int xh(x)\mu(dx) = \int xh^+(x)\mu(dx) - \int xh^-(x)\mu(dx)$ exists.) Then, (E,S_μ,H_μ) is ρ_μ-related. Finally, let H be any Hilbert space and $S : H \to E$ any continuous linear map such that (E,S,H) is ρ_μ-related. Then $I_\mu(x) = 1/2\|S^{-1}x\|_H^2$ for $x \in S(H)$ and $I_\mu(x) = \infty$ for $x \in E\setminus S(H)$. In particular, S is compact.

<u>Proof</u>: To see that (E,S_μ,H_μ) is ρ_μ-related, first note that $S_\mu^*(x^*) = x^*(\cdot)$ for $x^* \in E^*$. Indeed, for $y^* \in E^*$, $(y^*(\cdot),S_\mu^*(x^*))_{H_\mu} = x^*(\int xy^*(x)\mu(dx)) = \int x^*(x)y^*(x)\mu(dx) = (y^*(\cdot),x^*(\cdot))_{H_\mu}$. Since $\{y^*(\cdot) : y^* \in E^*\}$ is dense in H , this proves that $S_\mu^*(x^*) = x^*(\cdot)$.

In particular, $\|S_\mu^*(x^*)\|_{H_\mu}^2 = \|x^*(\cdot)\|_{H_\mu}^2 = \int |x^*(x)|^2 \mu(dx) = \rho_\mu(x^*,x^*)$. Also, if $h \in H$ and $S_\mu(h) = 0$, then for all $x^* \in E^*$: $0 = x_\mu^*(S_\mu(h)) = (h,S_\mu^*(x^*))_{H_\mu} = (h,x^*(\cdot))_{H_\mu}$. Since $\{x^*(\cdot) : x^* \in E^*\}$ is dense in H_μ , this proves that $S_\mu(h) = 0$ implies $h = 0$. Thus S_μ is injective. Since $\|S_\mu(h)\|_E = \|\int xh(x)\mu(dx)\|_E \leq (\int \|x\|_E^2 \mu(dx))^{1/2} (\int |h(x)|^2 \mu(dx))^{1/2} = (\int \|x\|_E^2 \mu(dx))^{1/2} \|h\|_{H_\mu}$, we have now proved that (E,S_μ,H_μ) is ρ_μ-related.

Now suppose that (E,S,H) is any ρ_μ-related triple. Noting that $\int e^{x^*(x)}\mu(dx) = e^{1/2\rho_\mu(x^*,x^*)}$, we see that $I_\mu(x) = \sup_{x^* \in E^*} (x^*(x) - 1/2\, \rho_\mu(x^*, x^*))$. Thus, by Lemma (3.44) , $I_\mu(x) = 1/2\|S^{-1}x\|_H^2$ if $x \in S(H)$ and $I_\mu(x) = \infty$ if $x \in E \setminus S(H)$. In particular, $\{S(h) : \|h\|_H \leq 1\} = \{x \in S(H) : \|S^{-1}(x)\|_H^2 \leq 1\} = \{x : I_\mu(x) \leq 1/2\}$ is compact. Thus, S is compact. □

(3.46) Exercise:

i) Let $E = C([0,T];R^d)$ and let μ denote Wiener measure on E (cf. (3.23)). Let $H = \{\phi \in E : \phi(0) = 0$ and $\dot\phi \in L^2([0,T];R^d)\}$ with the Hilbert norm $\|\phi\|_H = \|\dot\phi\|_{L^2([0,T];R^d)}$. If $S : H \to E$ is the inclusion map, show that (E,S,H) is ρ_μ-related and thereby give a second derivation of the result in (3.23) .

ii) Let $(\beta(t),\mathcal{F}_t,P)$ be a 1-dimensional Brownian motion starting at 0 and define $X(\cdot)$ as the solution to the integral equation

$$X(t) = \beta(t) - \int_0^t X(s)ds \quad , \quad 0 \leq t \leq 1 \quad . \tag{3.47}$$

Set $E = C([0,1];R^1)$ and let μ on E be the measure $P \circ X(\cdot)^{-1}$. Show that μ is a centered Gaussian and that $\rho_\mu(\alpha,\beta) = \int_0^1\int_0^1 k_\mu(s,t)\alpha(ds)\beta(dt)$ for $\alpha,\beta \in E^*$, where $k_\mu(s,t) = (e^{-|s-t|} - e^{-(s+t)})/2$. Finally, let

$H = L^2([0,1];R^1)$ with $\|\cdot\|_H = \|\cdot\|_{L^2([0,1];R^1)}$ and define $S : H \to E$ by $S(f)(t) = \int_0^t e^{-(t-s)} f(s)ds$, $0 \leq t \leq 1$. Show that (E,S,H) is ρ_μ-related, and deduce that $I(\phi) = \frac{1}{2}\int_0^1 |\dot{\phi}(t)+\phi(t)|^2 dt$ if $\phi(0) = 0$ and ϕ is absolutely continuous and that $I_\mu(\phi) = \infty$ otherwise.

We have twice seen how to prove Schilder's theorem as a consequence of our general theory. However, there is still one flaw in our derivations from the general theory, namely, our limits are along $\varepsilon = 1/n$ as $n \to \infty$. The next result removes this flaw.

(3.48) Theorem: Let μ be a centered Gaussian measure on the separable Banach space E . Then for any closed F ;

$$\overline{\lim_{\varepsilon \downarrow 0}}\, \varepsilon\mu\left(\frac{1}{\varepsilon^{1/2}} F\right) \leq \inf_{x \in F} I_\mu(x) \quad ,$$

and for any open G :

$$\underline{\lim_{\varepsilon \downarrow 0}}\, \varepsilon\mu\left(\frac{1}{\varepsilon^{1/2}} G\right) \geq -\inf_{x \in G} I_\mu(x) \quad .$$

Proof: First note that for any $n \geq 1$, $x/n^{1/2}$ is distributed under μ according to μ_n . Thus, if $0 < \varepsilon < 1$ and $n(\varepsilon) \geq 1$ is chosen so that $\frac{1}{n(\varepsilon)+1} \leq \varepsilon < \frac{1}{n(\varepsilon)}$, then $\varepsilon^{1/2}x$ is distributed under μ the same way as $\gamma(\varepsilon)^{1/2}x$ is under $\mu_{n(\varepsilon)}$, where $\gamma(\varepsilon) \equiv n(\varepsilon)\varepsilon \in [1-\varepsilon,1]$. Thus, if F is a closed set and $\tilde{F} = \{\frac{1}{\gamma^{1/2}} x : 1/2 \leq \gamma \leq 1$ and $x \in F\}$, then

$$\mu\left(\frac{1}{\varepsilon^{1/2}} F\right) = \mu_{n(\varepsilon)}\left(\frac{1}{\gamma(\varepsilon)^{1/2}} F\right) \leq \mu_{n(\varepsilon)}(\tilde{F})$$

for $0 < \varepsilon < 1/2$; and so

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log\mu\left(\frac{1}{\varepsilon^{1/2}}F\right) \leq \overline{\lim_{\varepsilon\downarrow 0}}\ \gamma(\varepsilon)\frac{1}{n(\varepsilon)}\log\mu_{n(\varepsilon)}(\tilde{F})$$

$$\leq -\inf_{x\in\tilde{F}} I_\mu(x) \quad .$$

At the same time, we see from Theorem (3.45) that

(3.49) $$I_\mu(\lambda x) = \lambda^2 I_\mu(x) \ , \quad \lambda \in R^1 \ \text{ and } \ x \in E \quad .$$

Thus

$$\inf_{x\in F} I_\mu(x) = \inf_{1/2\leq\gamma\leq 1}\ \inf_{x\in F} I_\mu\left(\frac{x}{\gamma^{1/2}}\right) = \inf_{1/2\leq\gamma\leq 1}\ \inf_{x\in F}\frac{1}{\gamma} I_\mu(x)$$

$$= \inf_{x\in F} I_\mu(x) \quad .$$

To complete the proof, let G be an open set and $x \in G$. Choose an open neighborhood U of x and an $\varepsilon_0 > 0$ so that $U \subseteq 1/\gamma(\varepsilon)^{1/2}\, G$ for all $0 < \varepsilon \leq \varepsilon_0$. Then

$$\underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon\log\mu\left(\frac{1}{\varepsilon^{1/2}}G\right) = \underline{\lim_{\varepsilon\downarrow 0}}\ \gamma(\varepsilon)\frac{1}{n(\varepsilon)}\log\mu_{n(\varepsilon)}\left(\frac{1}{\gamma(\varepsilon)^{1/2}}G\right)$$

$$\geq \underline{\lim_{\varepsilon\downarrow 0}}\ \frac{1}{n(\varepsilon)}\log\mu_{n(\varepsilon)}(U) \geq -I_\mu(x) \quad . \quad \square$$

(3.50 Corollary (Donsker-Varadhan): Let μ and E be as in Theorem (3.45) . Set $a = \inf\{I_\mu(x) : \|x\|_E = 1\}$. Then $a \in (0,\infty]$ and $a = \frac{1}{2b}$, where $b = \sup\{\rho_\mu(x^*,x^*) : \|x^*\|_E = 1\}$. Moreover,

$$\lim_{R\uparrow\infty}\frac{1}{R^2}\log\mu(\{x : \|x\|_E > R\}) = -a \quad .$$

In particular, $\int e^{\xi\|x\|_E^2}\mu(dx) < \infty$ for all $\xi \in (-\infty, a)$.

Proof: We first show that $\lim_{R\uparrow\infty} \frac{1}{R^2} \log \mu(\{x : \|x\|_E \geq R\}) = -a$. To this end, set $B = \{x : \|x\|_E < 1\}$. Then

$$\overline{\lim_{R\uparrow\infty}} \frac{1}{R^2} \log \mu(\{x : \|x\|_E \geq R\})$$

$$= \overline{\lim_{\varepsilon\downarrow 0}} \varepsilon \log \mu\left(\frac{1}{\varepsilon^{1/2}} B^c\right)$$

$$\leq -\inf_{x\in B^c} I_\mu(x) \quad ,$$

By (3.49) :

$$\inf_{x\in B^c} I_\mu(x) = \inf_{\lambda\geq 1} \inf_{\|x\|_E=1} \lambda^2 I_\mu(x) = \inf_{\|x\|_E=1} I_\mu(x) = a \quad .$$

At the same time:

$$\underline{\lim_{R\uparrow\infty}} \frac{1}{R^2} \log \mu(\{x : \|x\|_E > R\})$$

$$\geq \overline{\lim_{\varepsilon\downarrow 0}} \varepsilon \log \mu\left(\frac{1}{\varepsilon^{1/2}}(\overline{B})^c\right)$$

$$\geq -\inf_{\lambda>1} \inf_{\|x\|_E=1} \lambda^2 I_\mu(x) = -a \quad .$$

We must show that $a = \frac{1}{2b}$. To this end, let μ^{x^*} , $x^* \in E^*$, denote the distribution of $x^*(x)$ under μ . By Lemma (3.18) ,

$I_\mu(x) = \sup_{x^* \in E^*} I_{\mu^{x^*}}(x^*(x))$; and, by iii) of (3.10) , $I_{\mu^{x^*}}(x^*(x))$ $= \dfrac{(x^*(x))^2}{2\rho_\mu(x^*,x^*)}$. Thus, $I_\mu(x) = \sup_{x^* \in E^*} \dfrac{(x^*(x))^2}{2\rho_\mu(x^*,x^*)} = \sup\{\dfrac{(x^*(x))^2}{2\rho_\mu(x^*,x^*)} : \|x^*\|_E = 1\}$ $\geq \sup\{\dfrac{1}{2\rho_\mu(x^*,x^*)} : \|x^*\|_E = 1$ and $x^*(x) = 1\} \geq \dfrac{1}{2b}$ for all $\|x\|_E = 1$; and so $a \geq \dfrac{1}{2b}$. To prove the opposite inequality, let $x^* \in E^*$ with $\|x^*\|_{E^*} = 1$ be given. By the first part of this theorem and iii) of (3.10) :

$$\frac{-1}{2\rho_\mu(x^*,x^*)} = -I_{\mu^{x^*}}(1) = -\inf_{|\xi| \geq 1} I_{\mu^{x^*}}(\xi)$$

$$= \lim_{R\uparrow\infty} \frac{1}{R^2} \log \mu^{x^*}(\{\xi : |\xi| \geq R\})$$

$$\leq \lim_{R\uparrow\infty} \frac{1}{R^2} \log \mu(\{x : \|x\|_E \geq R\}) = -a \quad .$$

Thus $\dfrac{1}{2b} \geq a$.

We now have shown that $a = \dfrac{1}{2b}$. Since $b \leq \int \|x\|^2 \mu(dx) < \infty$, it follows that $a > 0$. □

It should not be surprising that everything is particularly elegant when E is a Hilbert space. The next result summarizes what can be said is this situation.

(3.51) Theorem: Let μ be a centered Gaussian on the separable Hilbert space E . Then there is a unique non-negative definite self-adjoint, trace class operator $R_\mu : E \to E$ satisfying $\rho_\mu(x,y) = (x,R_\mu y)_E$ for $x,y \in E$. Define $H = (\ker(R_\mu))^\perp$ and set $S = (R_\mu|_H)^{1/2}$. Then (E,H,S) is ρ_μ-related. In particular, if $R_\mu^{-1/2}$ is defined on $R_\mu^{1/2}(E)$ so that

$R_\mu^{-1/2}(x)$ is the unique $y \in H$ satisfying $x = R_\mu^{1/2}y$, then $I_\mu(x) = 1/2\|R_\mu^{-1/2}(x)\|_E^2$ if $x \in R_\mu^{1/2}(E)$ and $I_\mu(x) = \infty$ if $x \notin R_\mu^{1/2}(E)$.

Proof: Clearly there is a unique linear operator $R_\mu : E \to E$ defined by $(y,R_\mu x)_E = \rho_\mu(y,x)$ for all $x,y \in E$. Moreover, R_μ is obviously symmetric and non-negative. Also, if $\{e_n\}$ is an ortho-normal basis in E, then $\sum(e_n,R_\mu e_n)_E = \sum\rho_\mu(e_n,e_n) = \int \sum (e_n,x)_E^2\mu(dx) = \int\|x\|_E^2\mu(dx) < \infty$. Thus R_μ is trace class.

Next, note that if $x \in H$ and $S(x) = 0$, then $R_\mu(x) = R_\mu^{1/2} \circ S(x) = 0$ and so $x = 0$. Hence S is injective. Moreover, $S^* = R_\mu^{1/2}$ and so $\|S^*(x)\|_H^2 = \|R_\mu^{1/2}(x)\|_E^2 = (x,R_\mu x)_E = \rho_\mu(x,x)$. Hence (E,S,H) is ρ_μ-related; and so $I_\mu(x) = 1/2\|S^{-1}(x)\|_H^2 = 1/2\|R_\mu^{-1/2}(x)\|_E^2$ for $x \in S(H) = R_\mu^{1/2}(E)$ and $I_\mu(x) = \infty$ for $x \notin R_\mu^{1/2}(E)$. $\square$

Although in infinite dimensions the natural space on which a Gaussian measure lives is seldom a Hilbert space, it is often the case that the one can imbed the original space in a Hilbert space. The next simple lemma allows us to take advantage of such situations.

(3.52) Lemma: Let E_1 and E_2 be separable Banach spaces and let μ_1 be a probability measure on E_1 and $\Phi : E_1 \to E_2$ a continuous linear injection. Set $\mu_2 = \mu_1 \circ \Phi^{-1}$. Then $I_{\mu_1} = I_{\mu_2} \circ \Phi$.

Proof: Since Φ is an injection, $\Phi^*(E_2^*)$ is dense in E_1^*. Hence

$$I_{\mu_1}(x) = \sup_{x^*\in E_1^*} (x^*(x) - \log M_{\mu_1}(x^*))$$

$$= \sup_{y^* \in E_2^*} ((\Phi^*(y^*))(x) - \log M_{\mu_1}(\Phi(y^*)))$$

$$= \sup_{y^* \in E_2^*} (y^*(\Phi(x)) - \log M_{\mu_2}(y^*)) = I_{\mu_2} \circ \Phi(x) . \square$$

Now let K be a compact metric space and suppose μ is a centered Gaussian measure on $C(K;R^1)$. Thinking of $C(K;R^1)$ with the uniform norm, $C(K;R^1)$ becomes a separable Banach space. For $\xi,\eta \in K$, set $\rho_\mu(\xi,\eta) = \int x(\xi)x(\eta)\mu(dx)$. Then $\rho_\mu \in C(K \times K;R^1)$. In fact

$$\left|\rho_\mu(\xi_2,\eta_2) - \rho_\mu(\xi_1,\eta_1)\right| \leq \left|\rho_\mu(\xi_1,\eta_1) - \rho_\mu(\xi_2,\eta_1)\right| + \left|\rho_\mu(\xi_2,\eta_1) - \rho(\xi_2,\eta_2)\right| ,$$

and so, since ρ_μ is symmetric, we need only check that

$$\lim_{\xi_2 \to \xi_1} \sup_{\eta \in K} \left|\rho_\mu(\xi_2,\eta) - \rho_\mu(\xi_1,\eta)\right| = 0 .$$

But

$$\sup_{\eta \in K}\left|\rho_\mu(\xi_2,\eta) - \rho_\mu(\xi_1,\eta)\right| \leq \sup_{\eta \in K}\int\left|x(\xi_2) - x(\xi_1)\right|\left|x(\eta)\right|\mu(dx)$$

$$\leq \left(\int\left|x(\xi_2) - x(\xi_1)\right|^2\mu(dx)\right)^{1/2}\left(\int\|x\|^2_{C(K;R^1)}\mu(dx)\right)^{1/2}$$

Thus it suffices to check that

$$\lim_{\xi_2 \to \xi_1} \int\left|x(\xi_2) - x(\xi_1)\right|^2\mu(dx) = 0$$

However, $x(\xi_2) \to x(\xi_1)$ as $\xi_2 \to \xi_1$ for each $x \in C(K;R^1)$ and $\int \sup_{\xi_2 \in K} \left|x(\xi_2) - X(\xi_1)\right|^2\mu(dx) \leq 2\int\|x\|^2_{C(K;R^1)}\mu(dx) < \infty$. Thus, we are done.

We next note that if $\alpha,\beta \in C(K;R^1)^*$, then

$$\rho_\mu(\alpha,\beta) = \int_K\int_K \rho_\mu(\xi,\eta)\alpha(d\xi)\beta(d\eta) \quad . \tag{3.53}$$

In particular, let λ be a probability measure on K such that $\lambda(U) > 0$ for all open $U \neq \emptyset$. Let $\Phi : C(K;R^1) \to L^2(\lambda)$ be the natural imbedding and note that Φ is continuous and injective. Denote by $\tilde{\mu}$ the measure $\mu \circ \Phi^{-1}$ on $L^2(\lambda)$. From (3.53) , it is clear that

$$R_{\tilde{\mu}} f(\xi) = \int \rho_\mu(\xi,\eta) f(\eta)\lambda(d\eta) \quad , \quad f \in L^2(\lambda) \quad . \tag{3.54}$$

We can now prove the following.

(3.55) <u>Theorem:</u> If $\rho_\mu(\xi,\eta) = \int x(\xi)x(\eta)\mu(dx)$, $\xi,\eta \in K$, then $\rho_\mu \in C(K \times K;R^1)$ and

$$\lim_{R\uparrow\infty} \frac{1}{R^2} \log \mu(\{x : \|x\|_{C(K;R^1)} \geq R\}) = -\frac{1}{2b} \tag{3.56}$$

where $b = \sup_{\xi\in K} \rho_\mu(\xi,\xi)$. Moreover, if λ is a probabality measure on K which charges non-empty open sets and $R_{\tilde{\mu}}$ is defined by (3.54), then $R_{\tilde{\mu}}$ is a self-adjoint, trace class, non-negative operator on $L^2(\lambda)$ and $R_{\tilde{\mu}} f \in C(K;R^1)$ for all $f \in L^2(\lambda)$. Finally, if $R_\mu^{-1/2}$ is defined for $f \in R_{\tilde{\mu}}^{1/2}(L^2(\lambda))$ so that $R_\mu^{-1/2} f$ is the element g of $\ker(R_\mu)^\perp$ such that $R_\mu^{1/2} g = f$, then $I_\mu(x) = 1/2\|R_\mu^{-1/2}x\|^2_{L^2(\lambda)}$ for $x \in R_\mu^{1/2}(L^2(\lambda)) \cap C(K;R^1)$ and $I_\mu(x) = \infty$ otherwise.

<u>Proof:</u> Everything except (3.56) follows immediately from our preceding discussion. To prove (3.56) , note that by Schwartz's inequality

$$|\rho_\mu(\xi,\eta)|^2 \le \rho_\mu(\xi,\xi)\rho_\mu(\eta,\eta) \quad .$$

Thus, if $\alpha \in C(K;R^1)^*$ has total variation 1 , then

$$\rho_\mu(\alpha,\alpha) = \iint \rho_\mu(\xi,\eta)\alpha(d\xi)\alpha(d\eta)$$
$$\le \int \rho_\mu(\xi,\xi)|\alpha|(d\xi) \le b \quad .$$

On the other hand, if $\xi_0 \in K$ is chosen so that $\rho_\mu(\xi_0,\xi_0) = \sup_{\xi\in k} \rho_\mu(\xi,\xi)$ and if $\alpha_0 = \delta_{\xi_0}$, then $\rho_\mu(\alpha_0,\alpha_0) = b$. Thus, (3.56) follows from (3.50) . □

In order to be honest, it must be admitted that there is a far simpler proof that if μ is a centered Gaussian on a Banach space E , then there is an $\varepsilon > 0$ such that $\int e^{\varepsilon\|x\|_E^2}\mu(dx) < \infty$. The following argument is a beautiful discovery due to Fernique.

(3.57) Theorem: Let E be a separable Banach space and suppose that μ is a probability measure on E with the property that $(\frac{x_1+x_2}{2^{1/2}} , \frac{x_1-x_2}{2^{1/2}})$ is is distributed under μ^2 in the same way as (x_1,x_2) . Then, there is an $\varepsilon > 0$ such that $\int e^{\varepsilon\|x\|_E^2}\mu(dx) < \infty$. In particular, if μ is a centered Gaussian, then the conclusion holds.

Proof: If μ is a centered Gaussian, then an elementary covariance argument shows that $(\frac{x_1+x_2}{2^{1/2}} , \frac{x_1+x_2}{2^{1/2}})$ has the same distribution under μ^2 as does (x_1,x_2) . Thus we need only prove the first statement.

Now, let $0 < s < t$ be given. Then

$$\mu^2(\{(x_1,x_2) \in E^2 : \|x_1\|_E \le s , \|x_2\|_E \ge t\})$$
$$= \mu^2(\{(x_1,x_2) \in E^2 : \|x_1 - x_2\|_E \le 2^{1/2}s , \|x_1 + x_2\|_E \ge 2^{1/2}t\})$$

$$\leq \mu\ (\{(x_1,x_2) \in E^2 : \left|\|x_1\|_E - \|x_2\|_E\right| \leq 2^{1/2}s\ ,\ \|x_1\|_E + \|x_2\|_E \geq 2^{1/2}t\})$$

$$\leq \mu^2(\{(x_1,x_2) \in E^2 : \|x_1\|_E \wedge \|x_2\|_E \geq \frac{t-s}{2^{1/2}}\}) \quad ,$$

and so

$$\mu(\{x : \|x\|_E \leq s\})\mu(\{x : \|x\|_E \geq t\})$$

$$\leq (\mu(\{x : \|x\|_E \geq \frac{t-s}{2^{1/2}}\}))^2 \quad .$$

Taking $t_0 = s$ and $t_{n+1} = s + 2^{1/2}t_n$, we see that

$$\mu(\{x : \|x\| \geq t_{n+1}\})/\mu(\{x : \|x\| \leq s\}) \leq$$

$$(\mu(\{x : \|x\|_E \geq t_n\})/\mu(\{x : \|x\|_E \leq s\}))^2 \quad .$$

Hence: $\mu(\{x : \|x\|_E \geq t_n\}) \leq \mu(\{x : \|x\|_E \leq s\})\exp(2^n \log(\frac{\mu(\{x: \|x\|_E \geq s\})}{\mu(\{x: \|x\|_E \leq s\})}))$ for

for any $0 < s < t$ and $n \geq 0$. In particular, if s is chosen so that $\mu(\{x : \|x\|_E \geq s\})/\mu(\{x : \|x\|_E \leq s\}) = \rho < 1$, then:
$\mu(\{x : \|x\|_E \geq \frac{(2^{1/2})^{n+1}-1}{2^{1/2}-1}\ s) \leq e^{2^n\sigma}$, where $\sigma = \log\rho < 0$. From this, the desired result is clear. □

4. Large Deviation Principle for Diffusions:

In exercise (3.46) <u>ii)</u> we had an example of a diffusion, besides Brownian motion, for which we can derive a large deviations principle. Of course, our derivation in (3.46) <u>ii)</u> rested on the observation that if

$$X(t) = \beta(t) - \int_0^t X(s)ds \quad , \quad t \geq 0 \quad ,$$

where $\beta(\cdot)$ is a Brownian motion starting at 0 , then $X(\cdot)$ is a centered

Gaussian process. A second approach (and an approach which has a chance of generalizing to non-Gaussian diffusions) to obtaining a large deviations principle for $X(\cdot)$ is the following. Define $F : C([0,\infty);R^1) \to C([0,\infty);R^1)$ so that

$$(F(\psi))(t) = \psi(t) - \int_0^t (F(\psi))(s)ds \quad , \quad t \geq 0 \quad .$$

Clearly, for each $T > 0$, F determines a continuous injective surjection from $C([0,T];R^1)$ onto $C([0,T];R^1)$. Moreover, if $X_\varepsilon(\cdot) = \varepsilon^{1/2}X(\cdot)$, then $X_\varepsilon(\cdot) = F(\varepsilon^{1/2}\beta(\cdot))$. Hence, if μ denotes the distribution of $X(\cdot)\big|_{[0,T]}$, then for any closed $C \subseteq C([0,T];R^1)$:

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log \mu\Big(\frac{1}{\varepsilon^{1/2}} C\Big) = \overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P(\varepsilon^{1/2}\beta(\cdot) \in F^{-1}(C))$$

$$\leq -\inf\{I_{\mathcal{W}^T}(\psi) : F(\psi) \in C\} \quad ,$$

$I_{\mathcal{W}^T}$ is the rate function for Wiener measure $\mathcal{W}^T$ on $C([0,T];R^1)$. Because F is 1-1 and onto, we see that

$$\inf\{I_{\mathcal{W}^T}(\psi) : F(\psi) \in C\}$$

$$= \inf\{I_{\mathcal{W}^T} \circ F^{-1}(\psi) : \psi \in C\} \quad .$$

Finally, ψ is absolutely continuous if and only if $F(\psi)$ is; $\psi(0) = 0$ if and only if $(F(\psi))(0)$ does; and, if $\psi\big|_{[0,T]}$ is absolutely continuous, then $\dot\psi = (F(\psi))^{\bullet} + F(\psi)$. Therefore, if

$$I_\mu(\psi) = \begin{cases} 1/2\int_0^T \left|\dot\psi(t)+\psi(t)\right|^2 dt & \text{if } \psi(0) = 0 \text{ and } \psi\big|_{[0,T]} \text{ is absolutely continuous} \\ \infty & \text{otherwise,} \end{cases}$$

then

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log \mu\left(\frac{1}{\varepsilon^{1/2}} C\right) \leq -\inf_{\phi\in C} I_\mu(\phi) \quad .$$

Exactly the same reasoning leads to

$$\underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log \mu\left(\frac{1}{\varepsilon^{1/2}} G\right) \geq -\inf_{\phi\in G} I_\mu(\phi)$$

for open sets G .

To see that the above line of reasoning is not limited to Gaussian situations, let $b : R^1 \to R^1$ be any function satisfying $|b(y)-b(x)| \leq C|x-y|$, $x,y\in R^1$, for some $C < \infty$. For $\varepsilon > 0$, define $X_\varepsilon(\cdot)$ by

$$X_\varepsilon(t) = \varepsilon^{1/2}\beta(t) + \int_0^t b(X_\varepsilon(s))ds \quad , \quad t \geq 0 \quad .$$

Then, by precisely the same argument as we just used:

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P(X_\varepsilon(\cdot)\big|_{[0,T]} \in C) \leq -\inf_{\phi\in C} J(\phi)$$

for closed C in $C([0,T];R^1)$ and

$$\underline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P(X_\varepsilon(\cdot)\big|_{[0,T]} \in G) \geq -\inf_{\phi\in G} J(\phi)$$

for open G in $C([0,T];R^1)$, where

$$J(\phi) = \begin{cases} 1/2 \int_0^T |\dot\phi(t)-b(\phi(t))|^2 dt & \text{if } \phi(0) = 0 \text{ and } \phi\big|_{[0,T]} \text{ is absolutely continuous} \\ \infty & \text{otherwise.} \end{cases}$$

We now want to apply the preceding line of reasoning to more general situations. To be precise, we suppose that we are given:

(4.1) i) $\sigma : R^D \to R^D \otimes R^d$ and $b : R^D \to R^D$ satisfying

$$\|\sigma(y)-\sigma(x)\|_{H.S.(R^d;R^D)} \vee \|b(y)-b(y)\|_{R^D} \leq C|y-x|$$

for all $x,y \in R^D$ and some $C < \infty$;

ii) if $a = \sigma\sigma^*$, then there is an $\alpha > 0$ such that $a(x) \geq \alpha I_{R^D}$ for all $x \in R^D$

iii) there is an $M < \infty$ such that $\|\sigma(x)\|_{H.S.(R^d;R^D)} \vee |b(x)|_{R^D} \leq M$ for all $x \in R^D$.

Let $(\beta(t),\mathcal{F}_t,P)$ be a d-dimensional Brownian motion starting at 0 and for $x \in R^D$ and $\varepsilon > 0$, denote by $X^\varepsilon(\cdot,x)$ the unique solution to

$$(4.2)\qquad X^\varepsilon(T,x) = x + \varepsilon^{1/2}\int_0^T \sigma(X^\varepsilon(t,x))d\beta(t) + \int_0^T b(X^\varepsilon(t,x))dt \quad , \quad T \geq 0 \quad ;$$

and use P_x^ε on $C([0,\infty);R^D)$ to denote $P \circ X^\varepsilon(\cdot,x)^{-1}$ (the distribution of $X^\varepsilon(\cdot,x)$ under P). Clearly $P_x^\varepsilon \Rightarrow \delta_{\psi_0}$ as $\varepsilon \downarrow 0$, where $\psi_0(T) = x + \int_0^T b(\psi_0(t))dt$, $T \geq 0$. In order to prove a large deviation result, we would like to think of $X^\varepsilon(\cdot,x)$ as continuous function F of $\varepsilon^{1/2}\beta(\cdot)$ and proceed as in the preceding paragraph (where we had $a(\cdot) \equiv I$). Namely, we define F by:

$$\text{"}F(\psi)(T) = x + \int_0^T \sigma(F(\psi)(t))\dot{\psi}(t)dt + \int_0^T b(F(\psi))(t)dt,\text{"} \quad T \geq 0 \quad ;$$

and, just as before, we predict that the rate function $I_{x,T}^{a,b}$ for $\{P_x^\varepsilon : \varepsilon>0\}$ on $C([0,T];R^D)$ is given by

$$(4.3) \qquad I^{a,b}_{x,T}(\psi) = \begin{cases} 1/2\int_0^T (\dot\psi(t)-b(\psi(t)), a^{-1}(\psi(t))(\dot\psi(t)-b(\psi(t))))_{R^D}dt \\ \quad \text{if } \psi(0) = x \text{ and } \psi|_{[0,T]} \text{ is absolutely continuous,} \\ \infty \quad \text{otherwise.} \end{cases}$$

Unfortunately, we cannot "proceed as in the preceding paragraph" because the function F which we want to use is not continuous (or even well-defined, for that matter). Thus, we must find an appropriate mollification procedure in order to get around this technical difficulty.

Given $\varepsilon > 0$ and $n \geq 1$, define $X_n^\varepsilon(\cdot,x)$ to be the solution to:

$$(4.4) \qquad X_n^\varepsilon(T,x) = x + \varepsilon^{1/2}\int_0^T \sigma(X_n^\varepsilon([nt]/n,x))d\beta(t)$$

$$+ \int_0^T b(X_n^\varepsilon(t,x))dt \ , \quad T \geq 0 \ .$$

Then $X_n^\varepsilon(\cdot,x) = F_n(\varepsilon^{1/2}\beta(\cdot))$, where $F_n : C([0,\infty);R^D) \to C([0,\infty);R^D)$ is given by:

$$F_n(\psi)(0) = x$$

and, for $k \geq 0$:

$$F_n(\psi)(t) = F_n(\psi)(k/N) + \sigma(F_n(\psi)(k/n))(\psi(t)-\psi(k/n))$$

$$+ \int_{k/n}^t b(F_n(\psi)(s))ds \ , \quad k/n \leq t < \frac{k+1}{n} \ .$$

Note that for each $T > 0$, F_n is a continuous injective surjection from $\Omega_0(T) \equiv \{\psi \in C([0,T];R^D) : \psi(0) = 0\}$ onto $\Omega_x(T) \equiv \{\psi \in C([0,T];R^D): \psi(0) = x\}$. Hence, for each fixed $n \geq 1$ and $T > 0$, $\{P\circ(X_n^\varepsilon(\cdot,x)|_{[0,T]})^{-1} : \varepsilon > 0\}$ satisifies the large deviation principle with respect to:

$$(4.5)\qquad I_x^n(\phi) = \begin{cases} 1/2\int_0^T (\dot\phi(t)-b(\phi(t)),a^{-1}(\phi([nt]/n))(\dot\phi(t)-b(\phi(t))))_{R^D}dt \\ \quad \text{if } \phi(0) = x \text{ and } \phi|_{[0,T]} \text{ is absolutely continuous,} \\ \infty \quad \text{otherwise.} \end{cases}$$

(4.6) <u>Lemma</u>: For each $n \geq 1$, I_x^n is a rate function (i.e. $\{\phi \in C([0,T);R^D) : I_x^n(\phi) \leq L\}$ is compact for each $L > 0$). Also, $I_{x,T}^{a,b}$ is a rate function. Finally, for each closed set $C \subseteq C([0,T];R^D)$, $\inf_{\phi\in C} I_x^n(\phi) \to \inf_{\phi\in C} I_{x,T}^{a,b}(\phi)$ as $n \to \infty$.

<u>Proof</u>: A proof that I_x^n and $I_{x,T}^{a,b}$ are rate functions can be constructed along precisely the same lines as the proof of Lemma (1.8). Moreover, it is clear that for any $\phi \in C([0,T];R^D)$, $I_x^{a,b}(\phi) = \infty$ if and only if $I_x^n(\phi) = \infty$ for all $n \geq 1$. Thus, in proving that $\inf_{\phi\in C} I_x^n(\phi) \to \inf_{\phi\in C} I_{x,T}^{a,b}(\phi)$, we will assume that $\inf_{\phi\in C} I_{x,T}^{a,b}(\phi) < \infty$. But then (cf. Remark (2.12)) there is a $\phi_0 \in C$ such that $I_{x,T}^{a,b}(\phi_0) = \inf_{\phi\in C} I_{x,T}^{a,b}(\phi)$, and clearly $\overline{\lim}_{n\to\infty} \inf_{\phi\in C} I_x^n(\phi) \leq \overline{\lim}_{n\to\infty} I_x^n(\phi_0) = I_{x,T}^{a,b}(\phi_0)$. To complete the proof, note that if $K_L = \{\phi \in C([0,T];R^D) : \phi(0) = x \text{ and } \int_0^T |\dot\phi(t)|^2 dt \leq L\}$, then $\sup_{\phi\in K_L} |I_x^n(\phi) - I_{x,T}^{a,b}(\phi)| \to 0$. Since $\overline{\lim}_{n\to\infty} \inf_{\phi\in C} I_x^n(\phi) \leq \inf_{\phi\in C} I_{x,T}^{a,b}(\phi) < \infty$, we can choose $L < \infty$ so that $\inf_{\phi\in C} I_x^n(\phi) = \inf_{\phi\in C\cap K_L} I_x^n(\phi)$, $n \geq 1$, and $\inf_{\phi\in C} I_{x,T}^{a,b}(\phi) = \inf_{\phi\in C\cap K_L} I_{x,T}^{a,b}(\phi)$. Thus, choosing $\phi_n \in C \cap K_L$ so that $I_x^n(\phi_n) = \inf_{\phi\in C} I_x^n(\phi)$ and taking a subsequence $\{\phi_{n'}\}$ of $\{\phi_n\}$ so that

$I_x^{n'}(\psi_{n'}) \to \underline{\lim}_{n\to\infty} I_x^n(\psi_n)$ and $\psi_{n'} \to \phi$ (in $C([0,T];R^D)$), we conclude that:

$$\inf_{\psi\in C} I_{x,T}^{a,b}(\psi) \le I_{x,T}^{a,b}(\phi) \le \underline{\lim}_{n'\to\infty} I_{x,T}^{a,b}(\psi_{n'})$$

$$= \underline{\lim}_{n'\to\infty} I_x^{n'}(\psi_{n'}) = \underline{\lim}_{n\to\infty} \inf_{\psi\in C} I_x^n(\psi) \quad . \qquad \square$$

In order to complete the program, we must show that the $X_n^\varepsilon(\cdot,x) \to X^\varepsilon(\cdot,x)$ sufficiently fast that large deviation results for the $X_n^\varepsilon(\cdot,x)$'s can be transferred to $X^\varepsilon(\cdot,x)$.

(4.7) Lemma: Let $(\beta(t),\mathcal{F}_t,P)$ be a d-dimensional Brownian motion and suppose that $\alpha(\cdot)$ and $\gamma(\cdot)$ are $\mathcal{F}_t$ -progressively measurable with values in $R^D \otimes R^d$ and R^D , respectively. Assume that $\|\alpha(\cdot)\|_{H.S.} \le A < \infty$ and $|\gamma(\cdot)| \le B < \infty$, and set $\xi(T) = \int_0^T \alpha(t)d\beta(t) + \int_0^T \gamma(t)dt$, $T \ge 0$. Then, for $T > 0$ and $R > 0$ satisfying $D^{1/2}BT < R$:

$$P(\sup_{0\le t\le T} |\xi(t)| \ge R) \le 2D \exp(-(R-D^{1/2}BT)^2/2A^2DT) \quad .$$

Proof: The proof is similar to that of Lemma (1.11) . Namely, set $\overline{\xi}(t) = \xi(t) - \int_0^t \psi(s)ds$. Then , $\exp(\theta\cdot\overline{\xi}(t) - 1/2\int_0^t |\alpha^*(s)\theta|^2 ds)$ is a martingale for all $\theta \in R^D$. Hence, for fixed $\theta \in S^{D-1}$ and all $\lambda > 0$:

$$P(\sup_{0\le t\le T} \theta\cdot\xi(t) \ge R) \le P(\sup_{0\le t\le T} \theta\cdot\overline{\xi}(t) \ge R-BT)$$

$$\le P(\sup_{0\le t\le T} (\lambda\theta\cdot\overline{\xi}(t) - \frac{\lambda^2}{2}\int_0^t |\alpha^*(s)\theta|^2 ds) \ge \lambda(R-BT) - \frac{\lambda^2 A^2}{2} T)$$

$$\leq e^{-\lambda(R-BT)+\frac{\lambda^2A^2}{2}T} \quad .$$

Taking $\lambda = \frac{R-BT}{A^2T}$, we see that

$$P(\sup_{0\leq t\leq T} \theta\cdot\xi(t) \geq R) \leq \exp(-(R-BT)^2/2A^2T) \quad .$$

The rest of the proof is precisely like that of (1.11) . $\square$

(4.8) Lemma: For each $x \in R^D$, $T > 0$, and $\delta > 0$:

$$\lim_{n\to\infty} \overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P(\sup_{0\leq t\leq T}\left|X^\varepsilon(t,x) - X_n^\varepsilon(t,x)\right| \geq \delta) = -\infty \quad .$$

Proof: For convenience, we write $X^\varepsilon(\cdot)$ and $X_n^\varepsilon(\cdot)$ for $X^\varepsilon(\cdot,x)$ and $X_n^\varepsilon(\cdot,x)$, respectively. Set $Y_n^\varepsilon(\cdot) = X^\varepsilon(\cdot) - X_n^\varepsilon(\cdot)$. Then:

$$Y_n^\varepsilon(T) = \varepsilon^{1/2}\int_0^T(\sigma(X^\varepsilon(t)) - \sigma(X_n^\varepsilon[nt]/n))d\beta(t)$$

$$+ \int_0^T(b(X^\varepsilon(t)) - b(X_n^\varepsilon(t)))dt \quad , \quad T \geq 0 \quad .$$

For $\rho > 0$, define

$$\tau_{n,\rho}^\varepsilon = \inf\{t \geq 0 : \left|X_n^\varepsilon(t) - X_n^\varepsilon([nt]/n)\right| \geq \rho\} \quad ,$$

$$Y_{n,\rho}^\varepsilon(t) = Y_n^\varepsilon(t \wedge \tau_{n,\rho}^\varepsilon) \quad , \quad t \geq 0 \quad ,$$

and

$$\zeta_{n,\rho}^\varepsilon = \inf\{t \geq 0 : \left|Y_{n,\rho}^\varepsilon(t)\right| \geq \delta\} \quad .$$

Then

$$P(\sup_{0\le t\le T}|Y_n^\varepsilon(t)| \ge \delta) \le P(\tau_{n,\rho}^\varepsilon \le T) + P(\zeta_{n,\rho}^\varepsilon \le T) \quad . \tag{4.9}$$

First, note that

$$P(\tau_{n,\rho}^\varepsilon \le T) \le \sum_0^{[nT]} P(\sup_{k/n\le t\le\frac{k+1}{n}} |X_n^\varepsilon(t) - X_n^\varepsilon(k/n)| \ge \rho) \quad ;$$

and, by (4.7), for $n \ge 1$ satisfying $\frac{1}{n} D^{1/2}M \le \rho/2$,

$$P(\sup_{k/n\le t\le\frac{k+1}{n}} |X_n^\varepsilon(t) - X_n^\varepsilon(k/n)| \ge \rho)$$

$$\le 2D \exp(-n(\rho/2)^2/2DM^2\varepsilon) \quad .$$

Hence

$$\lim_{n\to\infty} \overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P(\tau_{n,\rho}^\varepsilon \le T) = -\infty \quad , \quad \rho > 0 \quad . \tag{4.10}$$

Now let $0 < \rho < 1$ be fixed and for $\lambda > 1$: set $\phi_\lambda(y) = (\rho^2 + |y|^2)^\lambda$, $y \in R^D$. By Itô's formula:

$$\phi_\lambda(Y_{n,\rho}^\varepsilon(t)) - \int_0^{t\wedge\tau_{n,\rho}^\varepsilon} \gamma_\lambda^\varepsilon(s)ds$$

is a martingale, where:

$$\begin{aligned}\gamma_\lambda^\varepsilon(t) &= 2\lambda(\rho^2 + |Y_n^\varepsilon(t)|^2)^{\lambda-1} Y_n^\varepsilon(t)\cdot(b(X^\varepsilon(t)) - b(X_n^\varepsilon(t))) \\ &+ 2\lambda(\lambda-1)\varepsilon|(\sigma(X^\varepsilon(t))-\sigma(X_n^\varepsilon([nt]/n)))^* Y_n^\varepsilon(t)|^2(\rho^2 + |Y_n^\varepsilon(t)|^2)^{\lambda-2} \\ &+ \lambda\varepsilon\|\sigma(X^\varepsilon(t))-\sigma(X_n^\varepsilon([nt]/n))\|_{H.S.}^2(\rho^2 + |Y_n^\varepsilon(t)|^2)^{\lambda-1} \quad .\end{aligned}$$

For $0 \leq t \leq \tau^{\varepsilon}_{n,\rho}$, $\left|\gamma_{\lambda}(t)\right| \leq C(\lambda(1+\varepsilon)+\lambda^{2}\varepsilon)\phi_{\lambda}(Y^{\varepsilon}_{n}(t))$ where $C < \infty$ is independent of ε, n, and λ . Taking $\lambda = 1/\varepsilon$ and setting $\Phi^{\varepsilon}_{n,\rho}(t) = E[(\rho^{2} + \left|Y^{\varepsilon}_{n,\rho}(t \quad \zeta^{\varepsilon}_{n,\rho})\right|^{2})^{1/\varepsilon}]$, we obtain

$$\Phi^{\varepsilon}_{n,\rho}(t) \leq \rho^{2/\varepsilon} + 3C/\varepsilon \int_{0}^{t} \Phi^{\varepsilon}_{n,\rho}(s)ds \quad , \quad t \geq 0 \quad ,$$

so long as $0 < \varepsilon \leq 1$. Hence

$$\Phi^{\varepsilon}_{n,\rho}(t) \leq \rho^{2/\varepsilon} e^{3Ct/\varepsilon} \quad .$$

But

$$(\rho^{2} + \delta^{2})^{1/\varepsilon} P(\zeta^{\varepsilon}_{n,\rho} \leq T) \leq \Phi^{\varepsilon}_{n,\rho}(T) \quad ,$$

and so

$$P(\zeta^{\varepsilon}_{n,\rho} \leq T) \leq \left(\frac{\rho^{2}}{\rho^{2}+\delta^{2}}\right)^{1/\varepsilon} e^{3CT/\varepsilon}$$

for $0 < \varepsilon \leq 1$. Thus

$$\overline{\lim_{\varepsilon \downarrow 0}}\ \varepsilon \log P(\zeta^{\varepsilon}_{n,\rho} \leq T) \leq \log \frac{\rho^{2}}{\rho^{2}+\delta^{2}} + 3CT \quad .$$

Finally, given $L > 0$, choose $0 < \rho < 1$ so that $\log \dfrac{\rho^{2}}{\rho^{2}+\delta^{2}} + 3CT \leq -2L$. Next, using (4.10) , choose N so that $\overline{\lim\limits_{\varepsilon \downarrow 0}}\ \varepsilon \log P(\tau^{\varepsilon}_{n,\rho} \leq T) \leq -2L$ for $n \geq N$. Then, for $n \geq N$ there is an $0 < \varepsilon_{n} < 1$ so that $P(\tau^{\varepsilon}_{n,\rho} \leq T) \leq e^{-L/\varepsilon}$ and $P(\zeta^{\varepsilon}_{n,\rho} \leq T) \leq e^{-L/\varepsilon}$ for $0 < \varepsilon \leq \varepsilon_{n}$; and so, by (4.9) :

$$P(\sup_{0 \leq t \leq T} \left|Y^{\varepsilon}_{n}(t)\right| \geq \delta) \leq 2e^{-L/\varepsilon}$$

for $0 < \varepsilon < \varepsilon_n$. Thus, $\overline{\lim_{\varepsilon \downarrow 0}} \varepsilon \log P(\sup_{0 \le t \le T} |Y_n^{\varepsilon}(t)| > \delta) \le -L$ for $n \ge N$. □

(4.11) Lemma: For each $T > 0$ and $x \in R^D$, $\{P_x^{\varepsilon}: \varepsilon > 0\}$ on $C([0,T];R^D)$ satisfies the large deviation principle with rate function $I_{x,T}^{a,b}$.

Proof: Let C be a closed subset of $C([0,T];R^D)$, and assume that $\emptyset \neq C \subseteq \{\psi : \psi(0) = x\}$. Given $\delta > 0$:

$$P(X^{\varepsilon}(\cdot,x) \in C) \le P(X_n^{\varepsilon}(\cdot,x) \in \overline{C^{\delta}})$$

$$+ P(\sup_{0 \le t \le T} |X^{\varepsilon}(t,x) - X_n^{\varepsilon}(t,x)| > \delta)$$

and so:

$$\overline{\lim_{\varepsilon \downarrow 0}} \varepsilon \log P(X^{\varepsilon}(\cdot,x) \in C)$$

$$\le [\overline{\lim_{\varepsilon \downarrow 0}} \varepsilon \log P(X_n^{\varepsilon}(\cdot,x) \in \overline{C^{\delta}})]$$

$$\vee [\overline{\lim_{\varepsilon \downarrow 0}} \varepsilon \log P(\sup_{0 \le t \le T} |X^{\varepsilon}(t,x) - X_n^{\varepsilon}(t,x)| \ge \delta)]$$

$$\le [-\inf_{\psi \in \overline{C^{\delta}}} I_x^n(\psi)] \vee [\overline{\lim_{\varepsilon \downarrow 0}} \varepsilon \log P(\sup_{0 \le t \le T} |X^{\varepsilon}(t,x) - X_n^{\varepsilon}(t,x)| \ge \delta)] .$$

Letting $n \to \infty$, we obtain:

$$\overline{\lim_{\varepsilon \downarrow 0}} \varepsilon \log P(X^{\varepsilon}(\cdot,x) \in C) \le -\inf_{\psi \in \overline{C^{\delta}}} I_{x,T}^{a,b}(\psi)$$

for all $\delta > 0$. At the same time (cf. Remark (2.4)), $\inf_{\psi \in \overline{C^{\delta}}} I_{x,T}^{a,b}(\psi) \to \inf_{\psi \in C} I_{x,T}^{a,b}(\psi)$. Thus the needed upper bound has been proved.

Next, let G be an open set in $C([0,T];R^D)$. Given $\psi^0 \in G$

satisfying $I^{a,b}_{x,T}(\psi^0) < \infty$, set $B_\rho = \{\psi : \sup_{0\le t\le T} |\psi(t)-\psi^0(t)| < \rho\}$ for $\rho > 0$. Choose $\delta > 0$ so that $B_{2\delta} \subseteq G$. Then

$$P(X^\varepsilon_n(\cdot,x) \in B_\delta) \le P(X^\varepsilon(\cdot,x) \in G)$$

$$+ P(\sup_{0\le t\le T} |X^\varepsilon(t,x)-X^\varepsilon_n(t,x)| \ge \delta) \quad .$$

Hence:

$$- I^n_x(\psi^0) \le \underline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P(X^\varepsilon_n(\cdot,x) \in B_\delta)$$

$$\le [\underline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P(X^\varepsilon(\cdot,x) \in G)]$$

$$\vee [\overline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P(\sup_{0\le t\le T} |X^\varepsilon(t,x) - X^\varepsilon_n(t,x)| \ge \delta)] \ .$$

Letting $n \to \infty$, we get:

$$\underline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P(X^\varepsilon(\cdot,x) \in G) \ge -I^{a,b}_{x,T}(\psi^0) \quad . \ \square$$

We have proved our large deviation result under the hypotheses listed in (4.1) . We will now show that condition iii) of (4.1) may be dropped.

(4.12) Lemma: Let $\sigma : R^D \to H.S.\ (R^d;R^D)$ and $b : R^D \to R^D$ be measurable functions satisfying $\|\sigma(y)\|_{H.S.} \le A(1+|y|)$ and $|b(y)| \le B(1+|y|)$, $y \in R^D$. Suppose that, for some $x \in R^D$,

$$X^\varepsilon(T) = x + \varepsilon^{1/2}\int_0^T \sigma(X^\varepsilon(t))d\beta(t) + \int_0^T b(X^\varepsilon(t))dt \quad , \quad T \ge 0 \ .$$

Then, for each $T > 0$,

$$\lim_{R\uparrow\infty} \overline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P(\sup_{0\le t\le T} |X^\varepsilon(t)| \ge R) = -\infty \ .$$

Proof: Without loss of generality, we assume that $x = 0$.

Set $\phi_\lambda(y) = (1+|y|^2)^\lambda$. Then, since $E[|X^\varepsilon(t)|^q] < \infty$ for all $\varepsilon > 0$, $t \geq 0$, and $q \in [1,\infty)$, we see by Itô's formula that:

$$\phi_\lambda(X^\varepsilon(t)) - \int_0^t \gamma_\lambda^\varepsilon(s)ds$$

is a martingale, where:

$$\begin{aligned}\gamma^\varepsilon(t) &= 2\lambda(1+|X^\varepsilon(t)|^2)^{\lambda-1}(b(X^\varepsilon(t)) , X^\varepsilon(t)) \\ &\quad + 2\lambda(\lambda-1)\varepsilon(1+|X^\varepsilon(t)|^2)^{\lambda-2}|\sigma^*(X^\varepsilon(t))X^\varepsilon(t)|^2 \\ &\quad + \lambda(1+|X^\varepsilon(t)|^2)^{\lambda-1}\varepsilon\|\sigma(X^\varepsilon(t))\|_{H.S.}^2 \\ &\leq C(\lambda + \lambda(\lambda+1)\varepsilon)\phi_\lambda(X^\varepsilon(t)) \quad .\end{aligned}$$

Thus, if $0 < \varepsilon \leq 1$ and we take $\zeta_R^\varepsilon = \inf\{t \geq 0 : |X^\varepsilon(t)| \geq R\}$, then:

$$E[(1+|X^\varepsilon(t\wedge\zeta_R^\varepsilon)|^2)^{1/\varepsilon}] \leq 1 + 3C/\varepsilon \int_0^t E[(1+|X^\varepsilon(s\wedge\zeta_R^\varepsilon)|^2)^{1/\varepsilon}]ds$$

and so:

$$E[(1+|X^\varepsilon(T\wedge\zeta_R^\varepsilon)|^2)^{1/\varepsilon}] \leq e^{3CT/\varepsilon} \quad .$$

In particular,

$$P(\zeta_R^\varepsilon \leq T) \leq (1+R^2)^{-1/\varepsilon}e^{3CT/\varepsilon} \quad ;$$

and therefore:

$$\overline{\lim_{\varepsilon\downarrow 0}}\, \varepsilon \log P(\sup_{0\leq t\leq T}|X^\varepsilon(t)| \geq R) \leq -\log(1+R^2) + 3CT \quad . \quad \square$$

(4.13) Theorem: Let $\sigma : R^D \to H.S.(R^d;R^D)$ and $b : R^D \to R^D$ be functions satisfying:

i) $\|\sigma(x)-\sigma(y)\|_{H.S.(R^d;R^D)} \vee |b(x)-b(y)| \le M|x-y|$ for all $x,y \in R^D$ and some $M < \infty$.

ii for each $x \in R^D$, $a(x) \equiv \sigma(x)\sigma^*(x) > 0$.

Define $I^{a,b}_{x,T}$ on $C([0,\infty);R^D)$ by (4.3). Then $I^{a,b}_{x,T}$ is a rate function. Moreover, if $X^\varepsilon(\cdot,x)$ is the solution to (4.2) and P^ε_x is the distribution of $X^\varepsilon(\cdot,x)$, then for each $x \in R^D$ and $T > 0$, $\{P^\varepsilon_x|_{\mathcal{M}_T} : \varepsilon > 0\}$ satisfies the large deviation principle with respect to $I^{a,b}_{x,T}$. (Here, $\mathcal{M}_T \equiv \sigma(x(t): 0 \le t \le T)$.)

Proof: Let $\eta \in C_0^\infty(R^D)$ be a non-negative function such that $\eta(y) = 1$ for $|y| \le 1$ and $\eta(y) = 0$ for $|y| \ge 2$. Given $R > 0$, set $\eta_R(y) = \eta(y/R)$ and define $a_R(y) = \eta_R(y)a(y) + (1-\eta_R(y))I_{R^D}$ and $b_R(y) = \eta_R(y)b(y)$. If $\sigma_R(y) = (a_R(y))^{1/2}$, then $\sigma_R(\cdot)$ and $b_R(\cdot)$ satisfy the conditions in (4.1) with $d = D$. Moreover, if $P^{\varepsilon,R}_x$ denotes the distribution of the solution to (4.2) with $\sigma = \sigma_R$ and $b = b_R$, then for any $A \in \mathcal{M}_T$, $P^{\varepsilon,R}_x(A \cap \{\zeta_R \ge T\}) = P^\varepsilon_x(A \cap \{\zeta_R \ge T\})$, where $\zeta_R = \inf\{t \ge 0 : |x(t)| \ge R\}$. Finally, $I^{a_R,b_R}_{x,T}(\phi) = I^{a,b}_{x,T}(\phi)$ for any $\phi \in C([0,\infty);R^D)$ satisfying $\|\phi\|^0_T \equiv \sup_{0\le t\le T}|\phi(t)| \le R$.

We now show that for any open $G \in \mathcal{M}_T$,

$$\underline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P^\varepsilon_x(G) \ge -\inf_{\phi\in G} I^{a,b}_{x,T}(\phi) . \tag{4.14}$$

Clearly, (4.14) will follow once we check that for all $\phi^0 \in C([0,\infty);R^D) \cap G$ satisfying $\phi^0(0) = x$:

$$\underline{\lim}_{\varepsilon\downarrow 0} \varepsilon \log P^\varepsilon_x(G) \ge -I^{a,b}_{x,T}(\phi^0) .$$

Given such a ϕ^0, choose $R > 0$ so that $\|\phi\|^0_T < R$ and set

$B_R = \{\phi : \|\phi\|_T^0 < R\}$. Then, by Lemma (4.11) :

$$\underline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log P_x^{\varepsilon}(G) \geq \underline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log P_x^{\varepsilon}(G \cap B_R)$$

$$= \underline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log P_x^{\varepsilon,R}(G \cap B_R)$$

$$\geq - \inf_{\phi \in G \cap B_R} I_{x,T}^{a_R,b_R}(\phi) \geq -I_{x,T}^{a,b}(\phi^0) \quad .$$

We next show that $I_{x,T}^{a,b}$ is a rate function. To this end, let $L < \infty$ be given. Using Lemma (4.12) , we can find an $R > 0$ so that:

$$\underline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log P_x^{\varepsilon}((\overline{B_R})^c) < -L \quad .$$

Thus, by (4.14)

$$\inf_{\phi \in (\overline{B_R})^c} I_{x,T}^{a,b}(\phi) > L \quad ,$$

and so $\{\phi : I_{x,T}^{a,b}(\phi) \leq L\} \subseteq \overline{B}_R$. But this means that:
$\{\phi : I_{x,T}^{a,b}(\phi) \leq L\} = \{\phi : I_{x,T}^{a,b}(\phi) \leq L\} \cap \overline{B}_R = \{\phi : I_{x,T}^{a_R,b_R}(\phi) \leq L\} \cap \overline{B}_R$, and the latter set is compact (in $C([0,T];R^D)$) because $I_{x,T}^{a_R,a_R}$ is a rate function.

Finally, given a closed C in $\mathcal{M}_T$ and an $L < \infty$, we can use (4.12) to find $R > 0$ so that

$$\overline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log P_x^{\varepsilon}(C)$$

$$\leq \underline{\lim}_{\varepsilon \downarrow 0} \varepsilon \log(P_x^{\varepsilon,R}(C \cap \overline{B}_R) + P_x^{\varepsilon}(B_R^c))$$

$$\leq \Big(- \inf_{\phi\in C\cap \overline{B}_R} I^{a_R, b_R}_{x,T}(\phi)\Big) \vee (-L)$$

$$\leq -[(\inf_{\phi\in C} I^{a,b}_{x,T}(\phi)) \wedge L] \quad .$$

Letting $L \uparrow \infty$, we obtain: $\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P^{\varepsilon}_{x}(C) \leq -\inf_{\phi\in C} I^{a,b}_{x,T}(\phi)$. □

In many applications, it is useful to have a "uniform version" of Theorem (4.13) . Our next result is such a version.

(4.14) <u>Theorem</u>: Under the assumptions in (4.13) about σ and b , if $\{x_{\varepsilon} : \varepsilon > 0\} \subseteq R^{D}$ and $x_{\varepsilon} \to x$ as $\varepsilon \downarrow 0$, then:

$$\underline{\lim}_{\varepsilon\downarrow 0}\ \varepsilon \log P^{\varepsilon}_{x_{\varepsilon}}(G) \geq -\inf_{\phi\in G} I^{a,b}_{x,T}(\phi)$$

for any open $\mathcal{M}_T$-measurable G , and

$$\overline{\lim_{\varepsilon\downarrow 0}}\ \varepsilon \log P^{\varepsilon}_{x_{\varepsilon}}(C) \leq -\inf_{\phi\in C} I^{a,b}_{x,T}(\phi)$$

for any closed $\mathcal{M}_T$-measurable C .

<u>Proof</u>: The proof turns on the estimate

(4.15) $$\lim_{R\uparrow\infty} \underline{\lim}_{\varepsilon\downarrow 0}\ \varepsilon \log P\Big(\sup_{0\leq t\leq T} \big|X^{\varepsilon}(t,x)-X^{\varepsilon}(t,y)\big| \geq R|x-y|\Big) = -\infty .$$

To derive (4.15) , set $Y^{\varepsilon}(t) = X^{\varepsilon}(t,x) - X^{\varepsilon}(t,y)$ and define $\zeta^{\varepsilon}_{R} = \inf\{t \geq 0 : |Y^{\varepsilon}(t)| \geq R|x-y|\}$. Applying Itô's formula to $\phi_{\varepsilon}(y) = |y|^{1/\varepsilon}$ (with $0 < \varepsilon \leq 1/2$) and proceeding as in the proof of Lemma (4.12) , we conclude that

$$E[|Y^{\varepsilon}(t\wedge\zeta^{\varepsilon}_{R})|^{1/\varepsilon}] \leq |x-y|^{1/\varepsilon} + C/\varepsilon \int_0^t E[|Y^{\varepsilon}(s\wedge\zeta^{\varepsilon}_{R})|^{1/\varepsilon}]ds .$$

Hence

$$P(\zeta_R^\varepsilon \le T) \le (1/R)^{1/\varepsilon} e^{CT/\varepsilon} \quad ,$$

and so (4.15) follows.

Given (4.15) , the rest of the argument parallels the derivation of Lemma (4.11) from Lemma (4.7) . The details are left as an exercise. $\square$

Following Ventcel and Freidlin, we now apply these results to study the way in which a randomly perturbed dynamical system exits from a region. In what follows, $\mathscr{G}$ will denote a bounded connected open set in R^D with smooth boundary $\partial\mathscr{G}$ and $\{P_x^\varepsilon : \varepsilon > 0 \text{ and } x \in R^D\}$ will be defined as in Theorem (4.13) . Define $\zeta = \inf\{t \ge 0 : x(t) \notin \mathscr{G}\}$.

(4.16) __Lemma__: For all $\varepsilon > 0$ and $x \in \mathscr{G}$, $E^{P_x^\varepsilon}[\zeta] < \infty$.

__Proof__: Given $x \in \mathscr{G}$, choose $R > 0$ so that $\mathscr{G} \subseteq B(x,R)$ and set $\tau_R = \inf\{t \ge 0 : |x(t)-x| \ge R\}$. Clear $\zeta \le \tau_R$. Now let $\theta \in S^{D-1}$ be fixed and set $\phi_\lambda(y) = e^{\lambda(\theta,y-x)}$ for $\lambda > 0$. By Itô's formula:

$$E^{P_x^\varepsilon}[\phi_\lambda(x(t\wedge\tau_R))] = 1 + E^{P_x^\varepsilon}[\int_0^{t\wedge\tau_R}[\frac{\varepsilon\lambda^2}{2}(\theta,a(x(s))\theta)+\lambda(\theta,b(x(s)))]\phi_\lambda(x(s))ds]$$

for all $t \ge 0$ and $\lambda > 0$. Noting that $(\theta,a(y)\theta) \ge \alpha > 0$, for all y and some $\alpha > 0$, and that $|b(y)| \le B < \infty$, for all $y \in \mathscr{G}$ and some $B < \infty$, we choose $\lambda > 0$ so that $\varepsilon\frac{\lambda^2}{2}\alpha - \lambda B \ge 1$ and thereby obtain:

$$E^{P_x^\varepsilon}[t \wedge \tau_R] \le e^{2\lambda R} , \quad t \ge 0 \quad . \square$$

Because of (4.16) , we can define the probability measure $\pi^\varepsilon(x,\cdot)$ on $\partial\mathscr{G}$ to be the distribution of $x(\zeta)$ under P_x^ε . Our goal is to study what

happens to $\pi^{\varepsilon}(x,\cdot)$ as $\varepsilon \downarrow 0$.

For each $x \in R^D$, denote by $\Phi(\cdot,x)$ the solution to

$$\Phi(T,x) = x + \int_0^T b(\Phi(t,x))dt \quad , \quad T \geq 0 \quad .$$

Clearly, $P_x^{\varepsilon} \Rightarrow \delta_{\Phi(\cdot,x)}$ as $\varepsilon \downarrow 0$. Suppose, for a moment, that $\Phi(\cdot,x)$ exits $\mathcal{G}$ in the sense that there is a $T > 0$ such that $\Phi(T,x) \notin \mathcal{G}$ and that if T_x is the first such T then $\Phi(t,x) \notin \mathcal{G}$ for all $t \in (T_x,T_x + \delta)$ for some $\delta > 0$ which is sufficiently small. Then, since $P_x^{\varepsilon}(\sup_{0\leq t\leq T_x+\delta} |x(t)-\Phi(t,x)| > \alpha) \to 0$ as $\varepsilon \downarrow 0$ for each $\alpha > 0$, we conclude that $\pi^{\varepsilon}(x,\cdot) \Rightarrow \delta_{\Phi(T_x,x)}$ as $\varepsilon \downarrow 0$. That is, nothing particularly interesting happens in this case.

We next consider the opposite situation. Namely, we suppose that:

(4.17)
<u>i)</u> for each $x \in \overline{\mathcal{G}}$, $\Phi(t,x) \in \mathcal{G}$ for all $t > 0$,
<u>ii)</u> there is an $x^0 \in \mathcal{G}$ such that for all $x \in \overline{\mathcal{G}}$, $\Phi(t,x) \to x^0$ as $t \to \infty$.

Now what happens to $\pi^{\varepsilon}(x,\cdot)$ as $\varepsilon \downarrow 0$? On the one hand, $P_x^{\varepsilon} \Rightarrow \delta_{\Phi(\cdot,x)}$. On the other hand, for each $\varepsilon > 0$, $\zeta < \infty$ (a.s., P_x^{ε}). Intuitively, what must be happening is that for any fixed $T > 0$, $x(\cdot)\big|_{[0,T]}$ under P^{ε} follows $\Phi(\cdot,x)\big|_{[0,T]}$ more and more closely as $\varepsilon \downarrow 0$ and then, after some large time T , $x(\cdot)$ abandons $\Phi(\cdot,x)$ and "makes a run for the boundary." When $x(\cdot)$ breaks and makes its "run," it must be quite close to x^0 . Thus we expect that during its "run" it should follow a route to $\partial\mathcal{G}$ which is an "efficient" route in the sense of $I^{a,b}_{x^0,T}$. With these considerations in mind, we now get down to business.

(4.18) <u>Lemma</u>: For every open neighborhood U of x^0 with $\overline{U} \subset\subset \mathcal{G}$ and

every $x \in \mathscr{S}$, $\lim\limits_{\varepsilon \downarrow 0} P_x^\varepsilon(\sigma_U < \zeta) = 1$, where $\sigma_U = \inf\{t \geq 0 : x(t) \in \overline{U}\}$.

Proof: Choose $T_x > 0$ so that $\Phi(T_x, x) \in U$ and choose $\delta > 0$ so that $B(\Phi(T_x,x),\delta) \subseteq U$ and $\mathrm{dist}(\Phi(t,x),\mathscr{S}) > \delta$ for all $0 \leq t \leq T_x$. Then $P_x^\varepsilon(\sigma_U \geq \zeta) \leq P_x^\varepsilon(\sup\limits_{0 \leq t \leq T_x} |x(t) - \Phi(t,x)| > \delta) \to 0$ as $\varepsilon \downarrow 0$. $\square$

For $x, y \in R^D$ and $T > 0$, define

$$I_T^{a,b}(x,y) = \inf\{I_{x,T}^{a,b}(\psi) : \psi(T) = y\}$$

and

$$I^{a,b}(x,y) = \inf_{T>0} I_T^{a,b}(x,y) \quad .$$

Note that for any $T > 0$, there is a ψ such that $I_T^{a,b}(x,y) = I_{x,T}^{a,b}(\psi)$. From this, it is clear that

$$I_{T_1+T_2}^{a,b}(x,y) \leq I_{T_1}^{a,b}(x,z) + I_{T_2}^{a,b}(z,y)$$

for all $T_1, T_2 > 0$ and $x,y,z \in R^D$. Hence

(4.19) $$I^{a,b}(x,y) \leq I^{a,b}(x,z) + I^{a,b}(z,y) \quad , \quad x,y,z \in R^D \quad .$$

In particular,

$$|I^{a,b}(x_1,y) - I^{a,b}(x_2,y)| \leq I^{a,b}(x_1,x_2) \vee I^{a,b}(x_2,x_1) \quad ,$$

$$|I^{a,b}(x,y_1) - I^{a,b}(x,y_2)| \leq I^{a,b}(y_1,y_2) \vee I^{a,b}(y_2,y_1) \quad ,$$

and so

(4.20) $$\left|I^{a,b}(x_1,y_1) - I^{a,b}(x_2,y_2)\right|$$

$$\leq I^{a,b}(x_1,x_2) \vee I^{a,b}(x_2,x_1) + I^{a,b}(y_1,y_2) \vee I^{a,b}(y_2,y_1) \quad .$$

(4.21) Lemma: For each $R > 0$, there is an $M_R < \infty$ such that $I^{a,b}(x,y) \leq M_R|x-y|$ for all $x,y \in \overline{B(0,R)}$. In particular, $I^{a,b} : \overline{B(0,R)} \times \overline{B(0,R)} \to R^1$ is uniformly Lipschitz continuous for each $R > 0$.

Proof: In view of (4.20), it suffices to prove the first statement. To do so, define $\phi(t) = x + \frac{y-x}{|y-x|} t$ for $0 \leq t \leq |y-x| \equiv T$. Clearly $I^{a,b}(x,y) \leq I^{a,b}_{x,T}(y) \leq M_R T$. □

Set $m = \inf_{y \in \partial \mathcal{D}} I^{a,b}(x^0,y)$ and $H = \{y \in \partial \mathcal{D}\colon I^{a,b}(x^0,y) = m\}$. Given $\delta > 0$, let $H(\delta) = \{y \in \partial \mathcal{D} : \text{dist}(y,H) < \delta\}$.

(4.22) Lemma: Let $\delta > 0$ be given. Then there is an open $U(\delta) \ni x^0$ satisfying $\overline{U(\delta)} \subset\subset \mathcal{D}$ and a $\gamma_\delta > 0$ such that

a) $\varliminf_{\varepsilon \downarrow 0} \inf_{x \in \overline{U(\delta)}} \varepsilon \log P^\varepsilon_x(x(\zeta) \in H(\delta),\ \zeta \leq T_\delta) \geq -m - \gamma_\delta$
for some $T_\delta > 0$,

b) $\varlimsup_{\varepsilon \downarrow 0} \sup_{x \in \overline{U(\delta)}} \varepsilon \log P^\varepsilon_x(x(\zeta) \notin H(\delta),\ \zeta \leq T) \leq -m - 2\gamma_\delta$
for all $T > 0$.

Proof: Choose $\alpha > 0$ so that $I^{a,b}(x^0,y) \geq m + \alpha$ for all $y \in \partial \mathcal{D} \backslash H(\delta)$. Next, choose $U(\delta) \ni x^0$ and $\rho_0 > 0$ so that: $\overline{U(\delta)} \subset\subset \mathcal{D}$; $I^{a,b}(x,y) \geq m + 3\alpha/4$ for all $x \in \overline{U(\delta)}$ and $y \in \partial \mathcal{D} \backslash H(\delta)$, and for each $x \in \overline{U(\delta)}$ there is a $\phi_x \in C([0,\infty);\mathcal{D})$ satisfying $\phi_x(0) = x$, $\phi_x(\rho_0) = x^0$, and $I^{a,b}_{x,\rho_0}(\phi_x) \leq \alpha/16$.

We first show that for any $T > 0$,

$\overline{\lim}_{\varepsilon\downarrow 0} \sup_{x\in\overline{U(\delta)}} \varepsilon \log P_x^\varepsilon(x(\zeta) \notin H(\delta) , \zeta \leq T) \leq -m - 3\alpha/4$. To this end, note that $\{x(\zeta) \notin H(\delta) , \zeta \leq T\}$ is a closed $\mathcal{M}_T$-measurable set. Suppose that there is a $\gamma > 0$ such that for $n \geq 1$ there is an x_n $\overline{U(\delta)}$ and an

$0 < \varepsilon_n \leq 1/n$ satisfying $\varepsilon_n \log P_{x_n}^{\varepsilon_n}(x(\zeta) \notin H(\delta) , \zeta \leq T) \geq -m - 3\alpha/4 + \gamma$.

Let $\{x_{n'}\}$ be a subsequence of $\{x_n\}$ such that $x_{n'} \to x \in \overline{U(\delta)}$. Then

$$\underline{\lim}_{n'\to\infty} \varepsilon_{n'} \log P_{x_{n'}}^{\varepsilon_{n'}}(x(\zeta) \notin H(\delta) , \zeta \leq T) \geq -m - 3\alpha/4 + \gamma .$$

At the same time, by Theorem (4.14) :

$$\begin{aligned}
&\overline{\lim}_{n'\to\infty} \varepsilon_{n'} \log P_{x_{n'}}^{\varepsilon_{n'}}(x(\zeta) \notin H(\delta) , \zeta \leq T) \\
&\leq -\inf\{I_{x,T}^{a,b}(\phi) : \zeta(\phi) \leq T \text{ and } \phi(\zeta(\phi)) \notin H(\delta)\} \\
&\leq -\inf\{I^{a,b}(x,y) : y \in \partial\mathcal{B}\setminus H(\delta)\} \\
&\leq -m - 3\alpha/4 .
\end{aligned}$$

Thus, our estimate is true.

In order to find a $T_\delta > 0$ for which $\underline{a)}$ holds, we proceed as follows. Choose $y^0 \in H$. Because $I^{a,b}(x^0,y^0) = m$, we can find a $T > 0$ and a $\phi_0 \in C([0,T];R^D)$ such that $\phi_0(0) = x^0$, $\phi_0(T) = y^0$, and $I_{x^0,T}^{a,b}(\phi\ ) \leq m + \alpha/16$. Define $T_0 = \inf\{t \geq 0 : \phi_0(t) \notin \mathcal{B}\}$. Clearly $T_0 \leq T$. Choose $\rho_1 > 0$ so that there is a $\phi_1 \in C([0,\rho_1];R^D)$ satisfying $\phi_1(0) = \phi_0(T_0)$, $\phi_1(t) \notin \overline{\mathcal{B}}$, for any $t \in (0,\rho_1]$, and $I_{\phi_0(T\ ),\rho_1}^{a,b}(\phi_1) \leq \alpha/16$

Set $T_\delta = \rho_0 + T_0 + \rho_1$ and for $x \in \overline{U(\delta)}$ define $\tilde{\Phi}_x \in C([0,T_\delta);R^D)$ by:

$$\tilde{\Phi}_x(t) = \begin{cases} \phi_x(t) , & 0 \leq t \leq \rho_0 \\ \phi_0(t-\rho_0) , & \rho_0 \leq t \leq T_0 + \rho_0 \\ \phi_1(t-T_0-\rho_0) , & T_0 + \rho_0 \leq t \leq T_\delta \end{cases} .$$

It is clear that

$$\sup_{x \in \overline{U(\delta)}} I^{a,b}_{x,T_\delta}(\phi_x) < m + \alpha/4 .$$

In particular, since

$$I^{a,b}(x,\phi_0(T_0)) \leq I^{a,b}_{x,T_0+\rho_0}(\tilde{\Phi}_x) \leq I^{a,b}_{x,T_\delta}(\tilde{\Phi}_x) ,$$

$\phi_0(T_0) \in H(\delta)$. Thus, for each $x \in \overline{U(\delta)}$, $\tilde{\Phi}_x$ is a path satisfying:

<u>i)</u> $\tilde{\Phi}_x(0) = x$, $T_0 + \rho_0$ is the first exit time of $\tilde{\Phi}_x$ from $\mathcal{D}$, $\tilde{\Phi}_x(T_0+\rho_0) \in H(\delta)$, and $\tilde{\Phi}(t) \notin \mathcal{D}$ for $t \in (T_0+\rho_0,T_\delta]$;

<u>ii)</u> $I^{a,b}_{x,T_\delta}(\tilde{\phi}_x) < m + \alpha/4$.

Now define $\overline{\zeta} = \inf\{t \leq 0 : x(t) \notin \overline{\mathcal{D}}\}$. Then $G \equiv \{x(\zeta) \in H(\delta) , \overline{\zeta} < T_\delta\}$ is an open $\mathcal{M}_{T_\delta}$-measurable set contained in $\{x(\zeta) \in H(\delta) , \zeta \leq T_\delta\}$ and $\tilde{\Phi}_x \in G$ for all $x \in \overline{U(\delta)}$. In particular,

$$\inf_{\phi \in G} I^{a,b}_{x,T_\delta}(\phi) \leq m + \alpha/4$$

for each $x \in \overline{U(\delta)}$. Thus, using Theorem (4.14) and arguing as we did before, we see that:

$$\varliminf_{\varepsilon \downarrow 0} \inf_{x \in \overline{U(\delta)}} \varepsilon \log P^{\varepsilon}_x(x(\zeta) \in H(\delta) , \zeta \leq T_\delta)$$

$$\geq \lim_{\varepsilon \downarrow 0} \inf_{x \in \overline{U(\delta)}} \varepsilon P_x^{\varepsilon}(G) \geq -m - \alpha/4 \quad . \qquad \Box$$

(4.23) Lemma: If U is an open neighborhood of x^0 with $\overline{U} \subset\subset \mathscr{G}$, then

$$\lim_{T \uparrow \infty} \inf\{I_{x,T}^{a,b}(\psi) : x \in \overline{\mathscr{G}} \setminus U \text{ and } \psi \in C([0,T]; \overline{\mathscr{G}} \setminus U)\} = \infty$$

Proof: Suppose not. Then there exists $M < \infty$, $\{x_n\} \subseteq \overline{\mathscr{G}} \setminus U$, $T_n \uparrow \infty$, and $\{\psi_n\} \subseteq C([0,T_n]; \overline{\mathscr{G}} \setminus U\}$ such that $I_{x_n,T_n}^{a,b}(\psi_n) \leq M$. By compactness, we may and will assume that $x_n \to x$ and that $\psi_n \to \psi$ uniformly on compacts. In particular, $\psi(t) \in \overline{\mathscr{G}} \setminus U$ for all $t \geq 0$ and $I_{x,T}^{a,b}(\psi) \leq M$ for all $T > 0$.

We next observe that there is a $T_0 > 0$ such that for each $x \in \overline{\mathscr{G}}$ $\Phi(t,x) \in U$ for some $t \in [0,T_0]$. Indeed, if $e_U(x) = \inf\{t \geq 0 : \Phi(t,x) \in U\}$, then $e_U(x) < \infty$ for all $x \in \overline{\mathscr{G}}$ and e_U is upper semi-continuous. Hence, since $\overline{\mathscr{G}}$ is compact, e_U is bounded on $\overline{\mathscr{G}}$

Now let ψ be the function produced in the first paragraph, and for $n \geq 0$ define $\psi_n \in C([0,T_0]; \overline{\mathscr{G}} \setminus U)$ by $\psi_n(t) = \psi(t + nT_0)$, $0 \leq t \leq T_0$. Then,

$$\sum_0^{\infty} I_{x_n,T_0}^{a,b}(\psi_n) = \lim_{T \uparrow \infty} I_{x,T}^{a,b}(\psi) \leq M < \infty$$

where $x_n = \psi(nT_0)$. Thus, there is a subsequence $\{\psi_{n'}\}$ such that $x_{n'} \to y$ and $I_{x_{n'},T_0}^{a,b}(\psi_{n'}) \to 0$. Clearly, $\psi_{n'} \to \Phi(\cdot,y)\big|_{[0,T_0]}$ in $C([0,T_0]; \overline{\mathscr{G}} \setminus U)$. In particular, $\Phi(t,y) \in \overline{\mathscr{G}} \setminus U$ for all $0 \leq t \leq T_0$; and this contradicts the choice of T_0. $\Box$

In order to complete our program we need to use the strong Markov property for the family $\{P_x^{\varepsilon} : x \in R^D\}$. Namely, given an $\{\mathcal{M}_t : t \geq 0\}$-stopping time $\tau : \Omega \to [0,\infty) \cup \{\infty\}$ and sets $A,B \in \mathcal{B}_{\Omega}$ with $A \in \mathcal{M}_t$:

$$P_x^{\varepsilon}(A \cap B^{\tau}) \tag{4.24}$$

$$= E^{P_x^{\varepsilon}}[P_{x(\tau)}^{\varepsilon}(B) \ , \ A \cap \{\tau < \infty\}] \ ,$$

where $B^{\tau} = \{\omega : \tau(\omega) < \infty$ and $x(\cdot + \tau(\omega),\omega) \in B\}$. A proof of (4.24) based on the same ideas as those used to prove Theorem (1.24) can be found in [S& V]. □

(4.25) Theorem: For every $x \in \mathscr{G}$ and $\delta > 0$, $\pi^{\varepsilon}(x,H(\delta)) \to 1$ as $\varepsilon \downarrow 0$. In particular, if H has precisely one element y^0 , then $\pi^{\varepsilon}(x,\cdot) \Rightarrow \delta_{y^0}$.

Proof: Given $\delta > 0$, choose $U(\delta)$ and T_{δ} as in Lemma (4.22) . Suppose that we can prove that

$$\lim_{\varepsilon \downarrow 0} \sup_{y \in \overline{U(\delta)}} P_y^{\varepsilon}(x(\zeta) \in \partial\mathscr{G}\backslash H(\delta)) = 0 \quad . \tag{4.26}$$

Then, by Lemma (4.18) and the strong Markov property:

$$\lim_{\varepsilon \downarrow 0} P_x^{\varepsilon}(x(\zeta) \in \partial\mathscr{G}\backslash H(\delta))$$

$$= \lim_{\varepsilon \downarrow 0} P_x^{\varepsilon}(x(\zeta) \in \partial\mathscr{G}\backslash H(\delta) \ , \ \sigma < \zeta)$$

$$= \lim_{\varepsilon \downarrow 0} E^{P_x^{\varepsilon}}[P_{x(\sigma)}^{\varepsilon}(x(\zeta) \in \partial\mathscr{G}\backslash H(\delta)) \ , \ \sigma < \zeta]$$

$$\leq \lim_{\varepsilon \downarrow 0} \sup_{y \in \overline{U(\delta)}} P_y^{\varepsilon}(x(\zeta) \in \partial\mathscr{G}\backslash H(\delta)) = 0 \quad ,$$

where $\sigma = \sigma_{U(\delta)} = \inf\{t \geq 0 : x(t) \in \overline{U(\delta)}\}$. Hence, it suffices to prove

(4.26) .

To prove (4.26) , let V be an open neighborhood of x^0 with $\overline{V} \subset\subset U(\delta)$. Define $\sigma_0 \equiv 0$,

$$\tau_n = \inf\{t \geq \sigma_n + T_\delta : x(t) \in \overline{V}\} \quad , \quad n \geq 0 \quad ,$$

and

$$\sigma_n = \inf\{t \geq \tau_{n-1} : x(t) \notin U(\delta)\} \quad , \quad n \geq 1 \quad .$$

Set

$$A_n = \{x(\zeta) \in H(\delta) \ , \ \zeta \in (\sigma_n, \tau_n)\}$$

and

$$B_n = \{x(\zeta) \in \partial\mathcal{S} \setminus H(\delta) \ , \ \zeta \in (\sigma_n, \tau_n)] \quad .$$

Then:

$$P_x^\varepsilon(x(\zeta) \in \partial\mathcal{S} \setminus H(\delta)) = \sum_0^\infty P_x^\varepsilon(B_n) \quad ,$$

$$P_x^\varepsilon(x(\zeta) \in H(\delta)) = \sum_0^\infty P_x^\varepsilon(A_n) \quad ,$$

and, for any $x \in \overline{U(\delta)}$,

$$P_x^\varepsilon(B_n) = E^{P_x^\varepsilon}[P_{x(\sigma_n)}^\varepsilon(x(\zeta) \in \partial\mathcal{S} \setminus H(\delta) \ , \ \zeta < \tau_0) \ , \ \sigma_n < \zeta]$$

$$\leq \theta(\varepsilon) E^{P_x^\varepsilon}[P_{x(\sigma_n)}^\varepsilon(x(\zeta) \in H(\delta) \ , \ \zeta < \tau_0) \ , \ \sigma_n < \zeta]$$

$$= \theta(\varepsilon) P_x^\varepsilon(A_n) \quad ,$$

where

$$\theta(\varepsilon) = \frac{\sup\limits_{y\in\overline{U(\delta)}} P_y^{\varepsilon}(x(\zeta) \in \partial\mathcal{G}\setminus H(\delta)\ ,\ \zeta < \tau_0)}{\inf\limits_{y\in\overline{U(\delta)}} P_y^{\varepsilon}(x(\zeta) \in H(\delta)\ ,\ \zeta < \tau_0)} \quad .$$

Thus:

$$\sup_{x\in\overline{U(\delta)}} P_x^{\varepsilon}(x(\zeta) \in \partial\mathcal{G}\setminus H(\delta)) \leq \theta(\varepsilon) \quad .$$

To prove that $\theta(\varepsilon) \to 0$ as $\varepsilon \downarrow 0$, first note that for any $T \geq T_\delta$:

$$\begin{aligned} & P_y^{\varepsilon}(x(\zeta) \in \partial\mathcal{G}\setminus H(\delta)\ ,\ \zeta < \tau_0) \\ & \leq P_y^{\varepsilon}(x(\zeta) \in \partial\mathcal{G}\setminus H(\delta)\ ,\ \zeta \leq T) \\ & \quad + P_y^{\varepsilon}(T < \zeta < \tau_0) \quad . \end{aligned}$$

Next, set $e_{\overline{V}} = \inf\{t \geq 0 : x(t) \in \overline{V}\}$. Then

$$\begin{aligned} P_y^{\varepsilon}(T < \zeta < \tau_0) &= E^{P_y^{\varepsilon}}[P^{\varepsilon}_{x(T_\delta)}(T-T_\delta < \zeta < e_{\overline{V}})\ ,\ \xi \wedge \tau_0 > T_\delta] \\ &\leq \sup_{z\in\mathcal{G}\setminus V} P_z^{\varepsilon}(x(t) \in \overline{\mathcal{G}}\setminus V \text{ for all } t \in [0,T-T_\delta]) \quad . \end{aligned}$$

By Lemma (4.23) , we can choose $T > T_\delta$ so that $\inf\{I^{a,b}_{z,T-T_\delta}(\phi) : z \in \overline{\mathcal{G}}\setminus V$ and $\phi \in C([0,T];\overline{\mathcal{G}\setminus V}\} \geq m + 2\gamma_\delta$. Hence, by Theorem (4.14) ,

$$\begin{aligned} \overline{\lim_{\varepsilon\downarrow 0}} \sup_{z\in\mathcal{G}\setminus V} \varepsilon \log P_z^{\varepsilon}(x(t) \in \overline{\mathcal{G}}\setminus V \text{ for all } t \in [0,T-T_\delta]) \\ \leq -m - 2\gamma_\delta \quad . \end{aligned}$$

At the same time, by (4.22) <u>b)</u> :

$$\overline{\lim_{\varepsilon\downarrow 0}} \sup_{y\in\overline{U(\delta)}} \varepsilon \log P_y^\varepsilon(x(\zeta) \in \partial\mathscr{G} \setminus H(\delta)\ ,\ \zeta \leq T) \leq -m - 2\gamma_\delta \ .$$

Hence, there is an $\varepsilon_0 > 0$ such that

$$\sup_{y\in\overline{U(\delta)}} P_y^\varepsilon(x(\zeta) \in \partial\mathscr{G}\setminus H(\delta)\ ,\ \zeta < \tau_0) \leq \exp(-(m+7\gamma_\delta/4)/\varepsilon)$$

for all $0 < \varepsilon \leq \varepsilon_0$. Finally,

$$P_y^\varepsilon(x(\zeta) \in H(\delta)\ ,\ \zeta < \tau_0) = P_y^\varepsilon(x(\zeta) \in H(\delta)\ ,\ \zeta \leq \tau_0)$$

$$\geq P_y^\varepsilon(x(\zeta) \in H(\delta)\ ,\ \zeta \leq T_\delta) \ .$$

Hence, by (4.22) a) , there is an $\varepsilon_1 > 0$ such that

$$\inf_{y\in\overline{U(\delta)}} P_y^\varepsilon(x(\zeta) \in H(\delta)\ ,\ \zeta < \tau_0)$$

$$\geq \exp(-(m + \frac{5\gamma_\delta}{4})/\varepsilon)$$

for $0 < \varepsilon \leq \varepsilon_1$. Therefore,

$$\theta(\varepsilon) \leq e^{-\gamma_\delta/2\varepsilon}$$

for $0 < \varepsilon \leq \varepsilon_0 \wedge \varepsilon_1$. □

5. Introduction to Large Deviations from Ergodic Phenomena:

Let E be a Polish metric space and set Ω equal to the space of maps $\omega : \{0,\dots,n,\dots\} \to E$. For each $n \geq 0$, define $X(n) : \Omega \to E$ so that $X(n,\omega)$ is the position of ω at time n and define $\theta_n : \Omega \to \Omega$ so that $X(m,\theta_n\omega) = X(m+n,\omega)$, $m \geq 0$.

Suppose that P on $(\Omega,\mathcal{B}_\Omega)$ is a $\theta_\bullet$ stationary ergodic probability

measure. That is, $P(\theta_m^{-1}A) = P(A)$ for all $m \geq 0$ and $A \in \mathcal{B}_\Omega$,and $P(A) \in \{0,1\}$ if $A = \theta_1^{-1}A$. Then, by the ergodic theorem, for each $\Phi \in L^1(P)$:

$$\frac{1}{N}\sum_0^{N-1} \Phi \circ \theta_n \to E^P[\Phi] \quad (\text{a.s.}, \ P)$$

as $N \uparrow \infty$. In particular, if $f : E \to R^1$ is bounded and measurable, then

$$(5.1) \qquad \frac{1}{N}\sum_0^{N-1} f(X(n)) \to \int_E f(y)\mu(dy) \quad (\text{a.s.}, \ P)$$

where $\mu = P \circ X(0)^{-1}$.

We now want to re-interpret (5.1) in such a way that it lends itself to the statement of a large deviation principle. To this end, denote by $\mathcal{M}_1(E)$ the space of probability measures on $(E,\mathcal{B}_E)$ and endow $\mathcal{M}_1(E)$ with the topology of weak convergence. For $N \geq 0$, define $L_N : \Omega \to \mathcal{M}_1(E)$ to be the "normalized occupation time" functional:

$$L_N(\Gamma,\omega) = \frac{1}{N}\sum_0^{N-1} \chi_\Gamma(X(n,\omega)) \quad .$$

Then (5.1) is equivalent to

$$(5.2) \qquad \int_E f(y)L_N(dy) \to \int_E f(y)\mu(dy) \quad (\text{a.s.}, \ P)$$

for all bounded measurable $f : E \to R^1$. Since E is Polish and therefore there exists a countable set $\mathcal{D} \subseteq C_b(E)$ such that $\nu_n \Rightarrow \nu$ if and only if $\int f d\nu_n \to \int f d\nu$ for each $f \in \mathcal{D}$, it follows from (5.2) that

$$(5.3) \qquad L_N \Rightarrow \mu \quad (\text{a.s.}, \ P) \quad .$$

In particular, if Q_N is defined on $\mathcal{M}_1(E)$ to be $P \circ L_N^{-1}$, then (5.3)

implies that

$$(5.4) \qquad Q_N \Rightarrow \delta_\mu$$

where the convergence here is weak convergence in $\mathcal{M}_1(\mathcal{M}_1(E))$.

The virtue of the formulation given in (5.4) is that it suggests how to formulate a large deviation principle. Namely, it suggests that we look at

$$\varliminf_{N\to\infty} \frac{1}{N} \log Q_N(G)$$

for open sets $G \subseteq \mathcal{M}_1(\mathcal{M}_1(E))$ and

$$\varlimsup_{N\to\infty} \frac{1}{N} \log Q_N(F)$$

for closed $F \subseteq \mathcal{M}_1(\mathcal{M}_1(E)$.

To see that it is not unreasonable to seek such a principle, we note that we have already proved one in a special case. Namely, suppose that $P = \alpha^{\mathcal{N}}$, where α is a probability measure on E and $\mathcal{N} = \{0,1,\ldots,n,\ldots\}$. Noting that the distribution of $\frac{1}{N}\sum_0^{N-1}\delta_{X(n)}$ under P is $P \circ L_N^{-1}$, we see that Sanov's theorem (Theorem (3.40)) ,provides us with precisely the kind of large deviation principle which we are after. In fact, if Λ_α is defined on $\mathcal{M}_1(E)$ as in (3.40) , then Q_N satisfies the large deviation principle with rate function Λ_α .

In order to understand what to expect when P is not simply a product measure, we consider a simple, but non-trivial, Markovian situation. Namely let E be a finite set and suppose that π is a transition probability function on E such that $\pi(x,\{y\}) > 0$ for all x,y E . Then, as an immediate consequence of the Frobenius theory of positive matrices, we know

that there is a unique π-stationary probability measure μ on E. In fact, $\mu(\{y\}) > 0$ for each $y \in E$ and there is an $\varepsilon > 0$ such that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \sup_{x\in E} \log\Big(\sum_{y\in E} \left|\pi^{(n)}(x,\{y\})-\mu(\{y\})\right|\Big) \leq -\varepsilon \quad .$$

Now let P on Ω be the Markov process with transition function π and initial distribution μ. Then P is a $\theta_{\bullet}$-stationary ergodic measure on Ω. We begin by trying to guess the rate function entering into the large deviation principle for the associated family $\{Q_n : n \geq 1\}$. In making our guess, we will assume that such a rate function I exists and will only attempt to see what form I has.

We begin by looking for an equation in which I appears. Such an equation is provided by Theorem (2.6):

$$\lim_{n\to\infty} \frac{1}{n} \log E^{Q_n}[e^{nF(\nu)}] = \sup_{\nu\in \mathcal{M}_1(E)} (F(\nu)-I(\nu))$$

for $F \in C_b(\mathcal{M}_1(E))$. In particular, if $V \in C_b(E)$, then

$$\lambda(V) \equiv \lim_{n\to\infty} \frac{1}{n} \log E^{Q_n}[e^{n\int_E V(y)\,\nu(dy)}] \tag{5.5}$$

$$= \sup_{\nu\in\mathcal{M}_1(E)} \left(\int V(y)\,\nu(dy)-I(\nu)\right) \quad .$$

Note that, since $\log E^{Q_n}[e^{n\int_E V(y)\,\nu(dy)}]$ is a convex function of $V \in C_b(E)$ for each $n \geq 1$, $V \to \lambda(V)$ is convex. Thus, if we assume that I is l.s.c. and convex on $\mathcal{M}_1(E)$ and we extend I to $C_b(E)^*$ so that $I(\nu) = \infty$ for $\nu \in C_b^*(E) \setminus \mathcal{M}_1(E)$, then (5.5) says that λ is the convex conjugate of I. In particular, this means that I is convex conjugate of λ (cf. Theorem 7.15). That is:

$$(5.6) \qquad I(\nu) = \sup_{V \in C_b(E)} \left(\int V(y)\nu(dy) - \lambda(V)\right) .$$

Before proceeding, we will check that (5.6) is precisely the formula from which Sanov's theorem came (cf. Lemma (3.37)). That is, suppose that $\pi(x,\cdot) = \alpha$ for all $x \in E$. Then $\mu = \alpha$ and $P = \alpha^{\eta}$. Hence $E^{Q_n}[e^{n\int_E V(y)\nu(dy)}] = E^P[\exp(\sum_0^{n-1} V(X(n)))] = E^P[\exp(V(X(0)))]^n = (\int e^{V(y)}\alpha(dy))^n$: and so $\lambda(V) = \log\int e^{V(y)}\alpha(dy)$. In other words, (5.6) becomes

$$I(\nu) = \sup_V \left(\int V(y)\nu(dy) - \log\int e^{V(y)}\alpha(dy)\right) ,$$

which is precisely the expression that appears in Lemma (3.37) and on which the proof of Sanov's theorem was based.

Returning to (5.6) , we next try to find a more tractable expression for $\lambda(V)$. To this end, we will prove the following.

(5.7) Lemma: Given $V \in C_b(E)$, define $\pi_V(x,\cdot) = e^{V(x)}\pi(x,\cdot)$. If P_x on Ω is the Markov process on Ω with transition function π starting from x , then

$$(5.8) \qquad [\pi_V^n f](x) = E^{P_x}[e^{\sum_0^{n-1} V(X(m))} f(X(n))] ,$$

where $\pi_V f(x) = \int f(y)\pi_V(x,dy)$, $\pi_V^{n+1} = \pi_V \circ \pi_V^n$, and $\sum_0^{n-1} V(X(m)) \equiv 0$ if $n = 0$.

Proof: We prove (5.8) by induction on $n \geq 0$. Clearly (5.8) holds when $n = 0$. Assuming (5.8) for n , we have:

$$\pi_V^{n+1} f(x) = e^{V(x)} E^{P_x}[\pi_V^n f(X(1))]$$

$$= e^{V(x)} E^{P_x}[E^{P_{X(1)}}[e^{\sum_0^{n-1} V(X(m))} f(X(n))]]$$

$$= e^{V(x)} E^{P_x}[e^{\sum_0^{n-1} V(X(m+1))} f(X(n+1))]$$

$$= E^{P_x}[e^{\sum_0^{n} V(X(m))} f(X(n+1))]$$

$\square$

In order to take advantage of (5.8) , we again invoke the Frobenius theorem and assert that:

i) There is a $\gamma(V) > 0$, a $\mu_V \in \mathcal{M}_1(E)$, and a $u_V \in C_b(E)$ such that $\int u_V(y)\mu_V(dy) = 1$, $u_V(y) \wedge \mu_V(\{y\}) > 0$ for all $y \in E$, $\pi_V u_V = \gamma(V)u_V$, and $\mu_V \pi_V = \gamma(V)\mu_V$.

ii) If $u \in C_b(E)^+ \setminus \{0\}$ and $\pi_V u = \gamma u$, then $\gamma = \gamma(V)$ and $u = \rho u_V$ where $\rho = \int u(y)\mu_V(dy)$.

iii) If $u \in C_b(E;C) \setminus \{0\}$ and $\gamma \in C$ satisfies $\pi_V u = \gamma u$, then either $u = \rho u_V$, where $\rho = \int u(y)\mu_V(dy)$, and $\gamma = \gamma(V)$ or $|\gamma| < \gamma(V)$.

iv) for all $f \in C_b(E)$, $(\frac{1}{\gamma(V)})^n \pi_V^n f \to (\int f(y)\mu_V(dy))u_V$ uniformly on E as $n \to \infty$.

Combining these facts with (5.8) , we see that

$$(\frac{1}{\gamma(V)})^n E^{Q_n}[e^{n\int_E V(y)\nu(dy)}] = (\frac{1}{\gamma(V)})^n E^P[e^{\sum_0^{n-1} V(X(m))}]$$

$$= \int_E (\frac{1}{\gamma(V)})^n [\pi_V^n 1](x)\mu(dx)$$

$$\to \int u_V(\chi)\mu(d\chi) > 0 \quad ;$$

and therefore that

$$\lambda(V) = \log \gamma(V) \tag{5.9}$$

where $\gamma(V) > 0$ is the largest eigenvalue of π_V . Next, from $\pi_V u_V = \gamma(V) u_V$, we see that $V - \lambda(V) = -\log \frac{\pi u_V}{u_V}$. Hence (5.6) can now be written as:

$$I(\nu) = -\inf_{V \in C_b(E)} \int_E \log\left(\frac{\pi u_V}{u_V}\right) d\nu \quad . \tag{5.10}$$

Finally, if $u \in C_b(E)$ is positive and $V_u = -\log\left(\frac{\pi u}{u}\right)$, then $\pi_{V_u} u = u$; and so $\gamma(V_u) = 1$ and $u = (\int u d\mu_{V_u}) u_{V_u}$. Hence every positive $u \in C_b(E)$ is a positive multiple of a u_V for some $V \in C_b(E)$. We can therefore replace (5.10) by:

$$I(\nu) = -\inf_{\substack{u \in C_b(E) \\ u > 0}} \int_E \log\left(\frac{\pi u}{u}\right) d\nu \quad . \tag{5.11}$$

Once again, notice that (5.11) is consistent with the expression for I_α in (3.37) ; indeed, simply take $u = e^f$ and remark that $\pi u = \int e^f d\alpha$.

Before ending this heuristic introduction, we look at the analogous set-up for continuous time processes. Again let E be a Polish space, but this time let $\Omega = D([0,\infty);E)$ be the space of right continuous maps $\omega : [0,\infty) \to E$ having left limits at every $t \in (0,\infty)$. Using Skorohod's topology, we can put a Polish metric on Ω in such a way that $\mathcal{B}_\Omega = \sigma(X(t) : t \geq 0)$, where $X(t) : \Omega \to E$ is defined so that $X(t,\omega)$ is the position of ω at time t . The time shift semi-group $\{\theta_t : t \geq 0\}$ is

defined on Ω so that $X(s,\theta_t\omega) = X(s+t,\omega)$, $s \geq 0$. Now suppose that P is a $\theta_\bullet$-stationary ergodic probability measure on Ω . Then, just as before, we see that if Q_t is defined on $\mathcal{M}_1(\mathcal{M}_1(E))$ by $Q_t = P \circ L_t^{-1}$ where

$$L_t(\Gamma,\omega) = \frac{1}{t}\int_0^t \chi_\Gamma(X(s,\omega))ds \quad , \quad \Gamma \in \mathcal{B}_E \quad ,$$

then the ergodic theorem leads to:

$$Q_t \Rightarrow \delta_\mu$$

as $t \uparrow \infty$. Thus, once again we are in a situation for which it is possible that a large deviation principle might hold.

In order to get a feeling for what to expect in this situation, we return to the case in which E is finite and we suppose that $(t,x) \in [0,\infty) \times E \to P(t,x,\bullet)$ is a continuous time, time homogeneous, Markovian transition function satisfying $P(t,x,\{y\}) > 0$ for all $(t,x,y) \in (0,\infty) \times E \times E$. Then there is a unique $\mu \in \mathcal{M}_1(E)$ satisfying $\mu = \mu P_t$ ($P_t f(x) = \int f(y)P(t,x,dy)$ and $\nu P_t(\Gamma) = \int P(t,x,\Gamma)\nu(dx)$) . Moreover, $P(t,x,\bullet) \to \mu$ as $t \uparrow \infty$ for all $x \in E$ and $\mu(\{y\}) > 0$ for all $y \in E$. Finally, if P on Ω is the Markov process with transition function $P(t,x,\bullet)$ and initial distribution μ , then P is a $\theta_\bullet$-stationary ergodic measure. What is the rate function I ?

As a first guess we proceed as follows. For each $h > 0$ let $\pi^{(h)}(x,\bullet) = P(h,x,\bullet)$ and define $L_n^{(h)}(\Gamma) = \frac{1}{n}\sum_0^{n-1} \chi_\Gamma(X(mh))$. By the preceding, we predict that the large deviation principle for $P \circ L_n^{(h)}$ has rate function $I^{(h)}(\nu) = -\inf\limits_{\substack{u \in C_b(E) \\ u > 0}} \int \log\frac{\pi^{(h)}u}{u}\, d\nu$. Noting that, for $f \in C_b(E)$:

$$h[t/h]\int f(y)L^{(h)}_{[t/h]}(dy) = h\sum_0^{[t/h]-1} f(X(mh))$$

$$\to t\int f(y)L_t(dy) \quad \text{as} \quad h \to 0 \ ,$$

we guess that:

$$\frac{1}{t} \log P(L_t \in A) \simeq \frac{1}{t} \log P(L^{(h)}_{[t/h]} \in A)$$

$$\simeq \frac{1}{h}\frac{1}{[t/h]} \log P(L^{(h)}_{[t/h]} \in A)$$

and therefore that

$$I(\nu) = -\inf_{\substack{u\in C_b(E)\\ u>0}} \lim_{h\downarrow 0} \frac{1}{h}\int \log \frac{\pi^{(h)}u}{u}\, d\nu$$

$$= -\inf_{\substack{u\in C_b(E)\\ u>0}} \frac{d}{dt}\int \log \frac{P_t u}{u}\, d\nu\Big|_{t=0} \ .$$

Hence, we guess that

$$I(\nu) = -\inf_{\substack{u\in C_b(E)\\ u>0}} \int \frac{Lu}{u}\, d\nu \tag{5.12}$$

where L is the generator of $\{P_t : t > 0\}$.

Although the preceding derivation of (5.12) is plausible, it involves several changes of the order in which limits are taken. We therefore will now give a second derivation, more in the spirit of the one used to guess (5.11) . To this end, note that, by Theorem (2.6) , for any $V \in C_b(E)$:

$$\lambda(V) \equiv \lim_{t\downarrow 0} \frac{1}{t} \log E^P[\exp(\int_0^t V(X(s))ds)]$$

$$= \lim_{t\downarrow 0} \frac{1}{t} \log E^{Q_t}[\exp(t\int_E V(y)\nu(dy))]$$

$$= \sup_{\nu \in \mathcal{M}_1(E)} \left(\int V(y)\nu(dy) - I(\nu) \right) .$$

Thus, if we assume that I is a l.s.c. convex function on $\mathcal{M}_1(E)$, then

$$(5.13) \qquad I(\nu) = \sup_{V \in C_b(E)} \left(\int V(y)\nu(dy) - \lambda(V) \right) .$$

We now need the analogue of Lemma (5.7) . To this end, let $\{P_t^V : t > 0\}$ be the semigroup on $C_b(E)$ generated by $L+V$. Then, by the standard perturbation theory, $\{P_t^V : t > 0\}$ is characterized as the unique solution to

$$(5.14) \qquad P_t^V f(x) = P_t f(x) + \int_0^t [P_{t-s}^V (VP_s f)](x) ds \quad , \quad t \geq 0$$

for $f \in C_b(E)$.

(5.15) Lemma (Feynman-Kac): For any $V \in C_b(E)$,

$$(5.16) \qquad P_t^V f(x) = E^{P_x}[\exp(\int_0^t V(X(s))ds) f(X(t))] \quad , \quad t \geq 0 \quad ,$$

for all $f \in C_b(E)$, where P_x is the Markov process on Ω with transition function $P(t,x,\cdot)$ starting at x .

Proof: Define $\{Q_t : t > 0\}$ by:

$$Q_t f(x) = E^{P_x}[\exp(\int_0^t V(X(s))ds) f(X(t))] \quad .$$

Then, by the Markov property:

$$Q_t f(x) - P_t f(x) = \int_0^t E^{P_x}[V(X(s))\exp(\int_0^s V(X(\sigma))d\sigma) f(X(t))] ds$$

$$= \int_0^t E^{P_x}[\exp(\int_0^s V(X(\sigma))d\sigma) V(X(s)) (P_{t-s} f)(X(s))] ds$$

$$= \int_0^t [Q_s(VP_{t-s}f)](x)ds$$

$$= \int_0^t [Q_{t-s}(VP_s f)](x)ds \quad .$$

That is, $\{Q_t : t > 0\}$ satisfies (5.14) and therefore $Q_t = P_t^V$. □

We once again invoke the Frobenius theory to conclude that there exists a unique $\alpha(V) \in R^1$, a unique positive $u_V \in C_b(E)$, and a unique $\mu_V \in \mathcal{M}_1(E)$, satisfying $\mu_V(\{y\}) > 0$ for all $y \in E$, such that $\int u_V(y)\mu_V(dy) = 1$ and $e^{-t\alpha(V)}P_t^V f \to (\int f(y)\mu_V(dy))u_V$ as $t \uparrow \infty$ for all $f \in C_b(E)$. Combining this with (5.16) , we now see that

$$\lambda(V) = \alpha(V) \quad .$$

Also, since

$$e^{-s\alpha(V)}P_s^V u_V = \lim_{t\uparrow\infty}(e^{-(s+t)\alpha(V)}P_{t+s}^V u_V) = u_V \quad ,$$

we see that

$$(L+V)u_V = \lambda(V)u_V \quad ;$$

and so

$$V - \lambda(V) = -Lu_V/u_V \quad .$$

Hence, we can replace (5.13) by

$$I(\nu) = -\inf_{V\in C_b(E)} \int_E (Lu_V/u_V)d\nu \quad .$$

Finally, just as in the discrete time case, if $u \in C_b(E)$ is positive and

$V_u = -Lu/u$, then $u = u_{V_u}$. Hence we again arrive at (5.12) .

In the case when L is symmetric on $L^2(\mu)$ there is a nicer way of computing $I(\nu)$ than the one given in (5.12) . To see this, we first observe that if $f,g \in C_b(E)$ then

$$(5.17) \qquad -2\int fLgd\mu = \lim_{t\downarrow 0} \frac{1}{t} \int [P_t((f-f(x))(g-g(x))](x)\mu(dx)$$

Indeed, define

$$\langle f,g\rangle_L = L(fg) - fLg - gLf$$

and note that

$$\begin{aligned}\int\langle f,g\rangle_L(x)\mu(dx) &= \int 1L(f\cdot g)d\mu - \int fLgd\mu - \int gLfd\mu \\ &= \int(f\cdot g)L1d\mu - 2\int fLgd\mu \\ &= -2\int fLgd\mu\end{aligned}$$

since $L1 = 0$. Next, note that P_t is symmetric in $L^2(\mu)$ and so

$$\begin{aligned}\int\langle f,g\rangle_L d\mu &= \lim_{t\downarrow 0} \frac{1}{t} \int([P_t(f\cdot g)](x) - f(x)[P_t g](x) \\ &\quad - g(x)[P_t f](x)+f(x)g(x))\mu(dx) \\ &= \lim_{t\downarrow 0} \frac{1}{t} \int[P_t((f-f(x))(g-g(x)))](x)\mu(dx) \quad .\end{aligned}$$

Now let $d\nu = \phi d\mu$ and apply (5.17) with $f = \phi/u$ and $g = u$. Then we see that

$$-2\int \frac{Lu}{u}\, d\nu = \lim_{t\downarrow 0} \frac{1}{t} \int[P_t((\phi/u-\phi/u(x))(u-u(x))](x)\mu(dx) \quad .$$

But:

$$[P_t((\phi/u - \phi/u(x))(u - u(x)))](x)$$

$$= \int (\phi(y) - (\frac{u(y)}{u(x)} \phi(x) + \frac{u(x)}{u(y)}\phi(y)) + \phi(x))P(t,x,dy)$$

$$\leq \int (f(y) - 2\phi^{1/2}(x)\phi^{1/2}(y) + \phi(x))P(t,x,dy)$$

$$= \int (\phi^{1/2}(y) - \phi^{1/2}(x))^2 P(t,x,dy) = [P_t(\phi^{1/2} - \phi^{1/2}(x))^2](x)$$

since $\lambda a + \frac{1}{\lambda} b \geq 2a^{1/2}b^{1/2}$ for all $a,b \geq 0$ and $\lambda > 0$. Hence, by (5.17):

$$-2\int \frac{Lu}{u} d\nu \leq \int [P_t(\phi^{1/2} - \phi^{1/2}(x))^2](x)d\mu$$

$$= -2\int \phi^{1/2} L \phi^{1/2} d\mu$$

This shows that

$$I(\nu) \leq -\int \phi^{1/2} L \phi^{1/2} d\mu \quad .$$

At the same time, if $u_n = \phi^{1/2} + 1/n$, then

$$I(\nu) \geq \overline{\lim_{n\to\infty}} - \int \frac{Lu_n}{u_n} d\nu$$

$$= \overline{\lim_{n\to\infty}} - \int \frac{\phi}{\phi^{1/2}+1/n} L\phi^{1/2} d\mu = -\int \phi^{1/2} L \phi^{1/2} d\mu \; .$$

Thus, we have now shown that when L is symmetric in $L^2(\mu)$, then

$$I(\nu) = -\int \phi^{1/2} L \phi^{1/2} d\mu \quad , \quad \phi = \frac{d\nu}{d\mu} \; . \tag{5.18}$$

It should be noted that (5.17) and (5.18) enable us to see that the equation from which we started, namely:

$$\lambda(V) = \sup_{\nu \in \mathcal{M}_1(E)} (\int V(y)\nu(dy) - I(\nu))$$

is precisely the classical variational principle for the largest eigenvalue

$\lambda(V)$ of the symmetric operator $L+V$. In fact, writing ν as $\phi^2\mu$, where $\phi \in L^2(\mu)^+$, we see that this formula is equivalent to:

$$\lambda(V) = \sup_{\substack{\phi\in L^2(\mu)^+ \\ \|\phi\|_{L^2(\mu)}=1}} \left(\int\phi(x)[(L+V)\phi](x)\mu(dx)\right)$$

Finally, the restriction of ϕ to $L^2(\mu)^+$ is inessential, since $\int V(x)\phi^2(x)\mu(dx) = \int V(x)\ |\phi(x)|^2\mu(dx)$ and, by (5.17) :

$$-2\int\phi L\phi d\mu = \lim_{t\downarrow 0}\frac{1}{t}\int[P_t(\phi-\phi(x))^2](x)\mu(dx)$$

$$= \lim_{t\downarrow 0}\frac{1}{t}\int\ \mu(dx)\int(\phi(y) - \phi(x))^2P(t,x,dy)$$

$$\geq \lim_{t\downarrow 0}\frac{1}{t}\int\mu(dx)\int(|\phi(y)|-|\phi(x)|)^2P(t,x,dy)$$

$$= -2\int\ |\phi|L(|\phi|)d\mu\ .$$

We have therefore shown that our variational formula for $\lambda(V)$ in terms of I is nothing but a hidden form of the classical formula

$$\lambda(V) = \sup_{\substack{\phi\in L^2(\mu) \\ \|\phi\|_{L^2(\mu)}=1}} \left(\int\phi(x)[(L+V)\phi](x)\mu(dx)\right)\ .$$

6. Existence of a Rate Function:

Let E be a Polish space and $\pi(x,\cdot)$ a transition probability function on E. Set $\Omega = E^{\eta}$ and let $\{P_x : x \in E\}$ be the Markov family on Ω with transition function π (i.e. $P_x(X(0) = x) = 1$ and P_x is a Markov process with transition function π). For $n \geq 1$ and $\omega \in \Omega$, define

$$L_n(\Gamma,\omega) = \frac{1}{n}\sum_0^{n-1} \chi_\Gamma(X(k,\omega)) \quad , \quad \Gamma \in \mathcal{B}_E \quad ,$$

and

$$L_n^1(\Gamma,\omega) = L_n^1(\Gamma,\theta_1\omega) = \frac{1}{n}\sum_1^n \chi_\Gamma(X(k,\omega)) \quad .$$

Finally, define $Q_{n,x}$ and $Q_{n,x}^1$ on $\mathcal{M}_1(\mathcal{M}_1(E))$ to be the distribution of L_n and L_n^1 , respectively, under P_x .

Under the assumption that for some $M < \infty$ and all $x_1, x_2 \in E$:

(6.1) $$\pi(x_1 \cdot) \leq M\pi(x_2 \cdot)$$

we are going to prove the following theorem.

(6.2) Theorem: There is a convex rate function $I : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ satisfying:

i) for each open G in $\mathcal{M}_1(E)$:

$$\varliminf_{n\to\infty} \frac{1}{n} \log \inf_{x\in E} Q_{n,x}(G) \geq -\inf_{\nu\in G} I(\nu) \quad ;$$

ii) for each closed F in $\mathcal{M}_1(E)$:

$$\varlimsup_{n\to\infty} \frac{1}{n} \log \sup_{x\in E} Q_{n,x}(F) \leq -\inf_{\nu\in F} I(\nu) \quad .$$

Moreover, if $\mu \in \mathcal{M}_1(E)$ satisfies $\varepsilon\mu \leq \pi(x_0,\cdot) \leq 1/\varepsilon\ \mu$ for some $\varepsilon > 0$ and $x_0 \in E$, then for any finite collection $\{A_1,\ldots,A_N\}$ of open convex sets A_i in $\mathcal{M}_1(E)$:

$$\lim_{n\to\infty} \frac{1}{n} \log Q_{n,\mu}(\bigcup_1^N A_i) = -\inf_{\nu\in\bigcup_1^N A_i} I(\nu) \quad ,$$

where $Q_{n,\mu} = \int Q_{n,x}\mu(dx)$.

We want to model our proof of Theorem (6.2) after our proof of Thorem (3.16) . The key to our ability to use the reasoning of section 3) in the present setting is contained in two simple remarks: we can replace $Q_{n,x}$ by $Q^1_{n,x}$ throughout most of the argument and

$$(6.3) \qquad \sup_x Q^1_{n,x} \leq M \inf_x Q^1_{n,x} \quad .$$

The first of these remarks will become clear in the final part of our argument. As for (6.3) , simply observe that for any $B \in \mathcal{B}_\Omega$ and all $x_1, x_2 \in \dot{E}$:

$$P_{x_1}(\theta_1^{-1}B) = \int_E P_y(B)\pi(x_1,dy)$$

$$\leq M\int_E P_y(B)\pi(x_2,dy) = MP_{x_2}(\theta_1^{-1}B) \quad .$$

(6.4) <u>Lemma</u>: For $A \in \mathcal{B}_{\mathcal{M}_1(E)}$, set $\mathcal{P}_n(A) = \inf_{x\in E} Q^1_{n,x}(A)$, $n \geq 1$. If A is open, then either $\mathcal{P}_n(A) = 0$ for all $n \geq 1$ or there is an $m \geq 1$ such that $\mathcal{P}_n(A) > 0$ whenever $n \geq m$. If A is convex, then $\mathcal{P}_{m+n}(A) \geq \mathcal{P}_m(A)\,\mathcal{P}_n(A)$ for all $m,n \geq 1$. Finally, if A is open and convex, then

$$\lim_{n\to\infty} \frac{1}{n} \log \mathcal{P}_n(A) = \sup_n \frac{1}{n} \log \mathcal{P}_n(A) \quad .$$

<u>Proof</u>: First suppose that A is convex. Then:

$$Q^1_{m+n,x}(A) = P_x(L^1_{m+n} \in A) = P_x(\tfrac{m}{m+n} L^1_m + \tfrac{n}{m+n} L^1_n \circ \theta_m \in A)$$

$$\geq P_x(L^1_m \in A \, , \, L^1_n \circ \theta_m \in A)$$

$$= \int_{\{L^1_m \in A\}} P_{X(m)}(L^1_n \in A)\,dP_x$$

$$\geq P_x(L^1_m \in A)\ \mathcal{P}_n(A) \geq \mathcal{P}_m(A)\ \mathcal{P}_n(A) \quad .$$

for all $x \in E$. Hence $\mathcal{P}_{m+n}(A) \geq \mathcal{P}_m(A)\ \mathcal{P}_n(A)$.

Next, suppose that A is open and that $\mathcal{P}_m(A) > 0$ for some $m \geq 1$. Select some $x_0 \in E$. Using Lemma (3.25) , choose a compact convex $K \subseteq A$ so that $Q^1_{m,x_0}(K) > 0$ and a $\delta > 0$ so that $\mu \in K$ and $\|\nu-\mu\|_{var} < 2\delta$ imply that $\nu \in A$. Given $n \geq [m/\delta] + 1$ and writing $n = q_n m + r_n$ with $0 \leq r_n < m$, we have:

$$Q^1_{n,x_0}(A) = P_{x_0}(L^1_n \in A)$$

$$\geq P_{x_0}(L^1_{q_n m} \in K)$$

$$= Q^1_{q_n m,x_0}(K)$$

$$\geq \mathcal{P}_{q_n m}(K) \geq (\ _m(K))^{q_n}$$

Since $\mathcal{P}_n(A) \geq \frac{1}{M} Q^1_{n,x_0}(A)$, this proves that $\mathcal{P}_n(A) > 0$ for all $n \geq [m/\delta] + 1$.

The final assertion is now an immediate corollary of Lemma (3.11) . □

Denote by $\overset{\circ}{\mathcal{C}}$ the class of open convex $A \subseteq \mathcal{M}_1(E)$ and by $\ell(A)$ the number $\sup_n \frac{1}{n} \log \mathcal{P}_n(A)$. For $\nu \in \mathcal{M}_1(E)$, define

$$I_\pi(\nu) = -\inf\{\ell(A) : \nu \in A \in \overset{\circ}{\mathcal{C}}\} \quad .$$

Before turning to the proof of (6.2) , we require one more observation.

(6.5) Lemma: For each $L > 0$ there is a compact K_L in $\mathcal{M}_1(E)$ such that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q^1_{n,x}(K_L^c)) < -L \quad .$$

Proof: Let $x_0 \in E$ be fixed and set $\mu = \pi(x_0,\cdot)$. For $n \geq 1$ and measurable $F : E^n \to [0,\infty)$, note that

$$\sup_{x\in E} E^{P_x}[F(X(1),\ldots X(n))] \leq M^n \int_{E^n} F(y)\mu^n(dy) \quad .$$

Indeed, this is obvious for $n = 1$ and follows by induction plus the Markov property for general $n \geq 1$.

Now, let $L > 0$ be given. By Lemma (3.32) , we can choose a compact K in $\mathcal{M}_1(E)$ so that $\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mu^n(\frac{1}{n}\sum_1^n \delta_{x_k} \in K^c) \leq -(L+\log M)$. Hence, from the preceding, we see that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q^1_{n,x}(K^c)) \leq -L$$

for this choice of K . □

(6.6) Proof of (6.2) with $I = I_\pi$:

Just as in Lemma (3.15) , lower semi-continuity and convexity of I_π are built into its definition in terms of $\mathcal{l}(A)$.

We next check the lower bound for open sets. Given an open G and $\nu \in G$, choose an $A \in \overset{\circ}{\mathfrak{C}}$ so that $\nu \in A$ and $\mathrm{dist.}(\overline{A},G^c) > 0$, where distance is measured in $\mathcal{M}_1(E)$ by the Lévy metric or any other convenient metric. Then, for all sufficiently large n's , $\inf_{x\in E} Q_{n,x}(G) \geq \mathcal{P}_n(A)$.

Hence,

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log(\inf_{x\in E} Q_{n,x}(G)) \geq \ell(A) \geq -I_\pi(\nu) \quad .$$

We have therefore proved that

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log(\inf_{x\in E} Q_{n,x}(G)) \geq -\inf_{\nu\in G} I_\pi(\nu) \quad .$$

Combining the preceding with Lemma (6.5), we see that for each $L > 0$ there is a compact K_L such that:

$$\begin{aligned} -\inf_{\nu\notin K_L} I_\pi(\nu) &\leq \underline{\lim_{n\to\infty}} \frac{1}{n} \log(\inf_{x\in E} Q_{n,x}(K_L^c)) \\ &\leq \underline{\lim_{n\to\infty}} \frac{1}{n} \log(\inf_{x\in E} Q^1_{n,x}(K^c)) \\ &\leq -L \quad , \end{aligned}$$

since $\inf_{x\in E} Q_{n,x}(A) \leq \inf_{x\in E} Q^1_{n,x}(A)$, $A \in \mathcal{B}_{\mathcal{M}_1(E)}$. In particular, $K_L^c \subseteq \{\nu : I_\pi(\nu) \geq L\}$; and so $\{\nu : I_\pi(\nu) \leq L\} \subseteq K_{L+1}$. Since we already know that I_π is l.s.c, this proves that $\{I_\pi \leq L\}$ is compact for every $L > 0$.

We now turn to the proof of the upper bound for closed sets F. To this end, note that it suffices to prove that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in F} Q^1_{n,x}(F)) \leq -\inf_{\nu\in F} I_\pi(\nu) \tag{6.7}$$

for all closed F. Indeed, for each $\delta > 0$,

$$\begin{aligned} &\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(F)) \\ &\leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(F^{(\delta)})) \quad , \end{aligned}$$

and so (6.7) would imply that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_x Q_{n,x}(F)) \leq -\inf_{\nu\in F^{(\delta)}} I_\pi(\nu)$$

for all $\delta > 0$. Since remark (2.4) applies to I_π , it is now clear that we need only check (6.7) .

Next note that because of (6.5) , it suffices to check (6.7) when F is compact. To see this, one simply repeats the argument given to prove Theorem (3.26) . Finally, if F is compact and $c > -\inf_{\nu\in F} I_\pi(\nu)$, choose $A_1,\ldots,A_N \in \overset{\circ}{\mathbf{C}}$ so that $c > \max_{1\leq i\leq N} \ell(A_i)$ and $F \subseteq \bigcup_1^N A_i$. Since $\sup_{x\in E} Q^1_{n,x}(F) \leq \sup_{x\in E} Q^1_{n,x}(\bigcup_1^N A_i) \leq MN \max_{1\leq i\leq N} \mathcal{P}_n(A_i)$ we then have:

$$\begin{aligned}
&\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q^1_{n,x}(F)) \\
&\leq \max_{1\leq i\leq N} \ell(A_i) < c \quad .
\end{aligned}$$

We have therefore proved the upper bound for all closed sets F . Note that, in particular, $0 = \overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q^1_{n,x}(\mathcal{M}_1(E))) \leq -\inf_{\nu\in\mathcal{M}_1(E)} I_\pi(\nu)$, and so $I_\pi \not\equiv \infty$. Thus we have also shown that I_π is a rate function.

To complete our proof, let $A_1,\ldots,A_N \in \overset{\circ}{\mathbf{C}}$ be given and set $A = \bigcup_1^N A_i$. Given $\mu \in \mathcal{M}_1(E)$ satisfying $\varepsilon\mu \leq \pi(x_0,\cdot) \leq 1/\varepsilon\, \mu$, note that

$$\begin{aligned}
\varepsilon Q^1_{n,x_0}(A) &= \varepsilon\int Q_{n,y}(A)\pi(x_0,dy) \\
&\leq \int Q_{n,y}(A)\mu(dy) \\
&\leq 1/\varepsilon\int Q_{n,y}(A)\pi(x_0,dy) = 1/\varepsilon\; Q^1_{n,x_0}(A) \quad .
\end{aligned}$$

Thus we need only check that $\lim_{n\to\infty} \frac{1}{n} \log Q^1_{n,x_0}(A) = -\inf_{\nu\in A} I_\pi(\nu)$. Moreover,

since

$$\max_{1\le i\le N} \mathcal{P}_n(A_i) \le Q^1_{n,x_0}(A) \le N \max_{1\le i\le N} Q^1_{n,x_0}(A_i)$$

$$\le MN \max_{1\le i\le N} \mathcal{P}_n(A_i)$$

and $\ell(A_i) = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{P}_n(A_i)$; we now see that it suffices to check that $\ell(B) = -\inf_{\nu\in B} I_\pi(\nu)$ for $B \in \overset{\circ}{\mathcal{C}}$; and clearly this reduces to showing that $\ell(B) \le -\inf_{\nu\in B} I_\pi(\nu)$ for $B \in \overset{\circ}{\mathcal{C}}$ satisfying $\ell(B) > -\infty$. Let such a B be given. For a fixed $\delta > 0$, choose $N_0 \ge 1$ so that $\frac{1}{N_0} \log M < \delta$ and $\ell(B) - \frac{1}{N_0} \log Q^1_{N_0,x_0}(B) < \delta$. Next, using Lemma (3.25) , choose a compact convex $K \subseteq B$ so that $\frac{1}{N_0} \log Q^1_{N_0,x_0}(B) - \frac{1}{N} \log Q^1_{N_0,x_0}(K) < \delta$. Then

$$\frac{1}{N_0}\log \mathcal{P}_{N_0}(K) \ge \frac{1}{N_0}\log Q^1_{N_0,x_0}(K) - \frac{1}{N_0}\log M$$

$$\ge \ell(B) - 3\delta \quad .$$

Since $n \to \log \mathcal{P}_n(K)$ is super-additive, this proves that $\underline{\lim}_{n\to\infty} \frac{1}{n} \log \mathcal{P}_n(K) \ge \ell(B) - 3\delta$. At the same time,

$$\mathcal{P}_n(K) \le Q^1_{n,x_0}(K) \le \sup_{x\in E} Q_{n,x}(K)$$

and so

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathcal{P}_n(K) \le \overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(K))$$

$$\le -\inf_{\nu\in K} I_\pi(\nu) \le -\inf_{\nu\in B} I_\pi(\nu) \quad .$$

□

We next turn to the continuous time context. Again let E be a Polish space and this time set Ω equal to the Polish space $D([0,\infty);E)$ of right-continuous maps $\omega : [0,\infty) \to E$ having a left limit at each $t \in (0,\infty)$.

Suppose that $\{P_x : x \in E\}$ is a measurable, time homogeneous Markov family of probability measures on $(\Omega, \mathcal{B}_\Omega)$, and denote by $P(t,x,\cdot)$ its transition probability function. For each $t > 0$, define $L_t : \Omega \to \mathcal{M}_1(E)$ so that

$$L_t(\Gamma,\omega) = \frac{1}{t}\int_0^t \chi_\Gamma(X(s,\omega))ds \quad , \quad \Gamma \in \mathcal{B}_E \quad ,$$

and set

$$L_t^1(\Gamma,\omega) = L_t(\Gamma,\theta_1\ \omega) = \frac{1}{t}\int_1^{t+1} \chi_\Gamma(X(s,\omega))ds \quad .$$

Let $Q_{t,x} = P_x \circ L_t^{-1}$ and $Q_{t,x}^1 = P_x \circ (L_t^1)^{-1}$ on $(\mathcal{M}_1(E), \mathcal{B}_{\mathcal{M}_1(E)})$. Under the assumption that

$$(6.8) \qquad P(1,x_1,\cdot) \leq MP(1,x_2,\cdot)$$

for some $M < \infty$ and all $x_1, x_2 \in E$, we are going to prove the following theorem.

(6.9) Theorem: There is a convex rate function $I : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ satisfying:

i) for all open G in $\mathcal{M}_1(E)$:

$$\varliminf_{t\to\infty} \frac{1}{t} \log(\inf_{x\in E} Q_{t,x}(G)) \geq -\inf_{\nu\in G} I(\nu) \quad ;$$

ii) for all closed F in $\mathcal{M}_1(E)$:

$$\varlimsup_{t\to\infty} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(F)) \leq -\inf_{\nu\in F} I(\nu) \quad .$$

The proof of Theorem (6.9) is somewhat a re-run of the discrete case. However, a few alterations are necessary.

Given $A \in \mathcal{B}_{\mathcal{M}_1(E)}$, define

$$\mathcal{P}_t(A) = \inf_{x \in E} Q_{t,x}(A)$$

for $t > 0$.

(6.10) Lemma: For any convex $A \in \mathcal{B}_{\mathcal{M}_1(E)}$, $\mathcal{P}_{s+t}(A) \geq \mathcal{P}_s(A)\mathcal{P}_t(A)$, $s,t > 0$. Moreover, if $A \in \overset{\circ}{C}$ and there exists a $B \in \overset{\circ}{C}$ satisfying $(\mathrm{dist}(B,A^c)) \wedge (\sup_{t>0} \mathcal{P}_t(B)) > 0$, then $\lim_{t\to\infty} \frac{1}{t} \log \mathcal{P}_t(A) = \sup_{t>0} \frac{1}{t} \log \mathcal{P}_t(A)$.

Proof: The first part follows from:

$$Q_{s+t,x}(A) = P_x\left(\frac{s}{s+t} L_s + \frac{t}{s+t} L_t \circ \theta_s \in A\right) \geq P_x(L_s \in A\ ,\ L_t \circ \theta_s \in A)$$

$$= \int_{\{L_s \in A\}} P_{X(s)}(L_t \in A)dP_x \geq \mathcal{P}_t(A)P_x(L_s \in A) .$$

In view of the first part and exercise (6.11) below, the second part will follow as soon as we show that there is a non-empty open interval $J \subseteq (0,\infty)$ with the property that $\inf_{t \in J} \mathcal{P}_t(A) > 0$. To this end, first choose $T_0 > 0$ so that $Q^1_{t,x}(B) \leq Q_{t,x}(A)$ for all $t \geq T_0$ and $x \in E$. Next, note that since $\sup_{t>0} \mathcal{P}_t(B) > 0$ and $t \to \mathcal{P}_t(B)$ is super-multiplicative, there is a $T \geq T_0$ such that $\mathcal{P}_T(B) > 0$. Now fix $x_0 \in E$ and note that $Q^1_{T,x_0}(B) \geq \mathcal{P}_T(B) > 0$. Since $t \to L_t$ is continuous (a.s., P_{x_0}), it is clear that $t \to Q^1_{t,x_0}(B)$ is l.s.c. Hence there is a $\delta > 0$ and an $\varepsilon > 0$ such that $Q^1_{t,x_0}(B) \geq \varepsilon$ for $t \in [T,T+\delta]$. At the same time,

$$Q^1_{t,x_0}(B) = \int Q_{t,y}(B)P(1,x_0,dy) \leq M\int Q_{t,y}(B)P(1,x,dy)$$

$$= MQ^1_{t,x}(B) .$$

Thus $Q^1_{t,x}(B) \geq \varepsilon/M$ for all $t \in [T,T+\delta]$ and $x \in E$. Finally, since $Q_{t,x}(A) \geq Q^1_{t,x}(B)$ for all $t \geq T_0$ and $x \in E$, we conclude that $\mathcal{P}_t(A) \geq \varepsilon/M$ for all $t \in [T,T+\delta]$. □

(6.11) <u>Exercise</u>: Let $f : (0,\infty) \to [0,\infty) \cup \{\infty\}$ be sub-additive (i.e. $f(s+t) \leq f(s) + f(t)$ for all $s,t > 0$). Further, assume that there is a non-empty open interval $J \subseteq (0,\infty)$ such that $\sup_{t\in J} f(t) < \infty$. Show that $\lim_{t\to\infty} \frac{1}{t} f(t) = \inf_{t>0} \frac{1}{t} f(t)$. An argument can be modeled on the proof of Lemma (3.11) .

We now define $\ell(A)$ for $A \in \mathring{\mathbf{C}}$ as follows:

$$\ell(A) = \begin{cases} \sup_{t>0} \frac{1}{t} \log \mathcal{P}_t(A) & \text{if } (\mathrm{dist}(\overline{B},A^c)) \wedge (\sup_{t>0} \mathcal{P}_t(B)) > 0 \text{ for some } B \in \mathring{\mathbf{C}} \\ -\infty & \text{otherwise.} \end{cases}$$

Notice that $\ell(B) \leq \ell(A)$ if $A,B \in \mathring{\mathbf{C}}$ and $B \subseteq A$. Define $I_p : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ by:

$$I_p(\nu) = -\inf\{\ell(A) : \nu \in A \in \mathring{\mathbf{C}}\} .$$

(6.12) <u>Lemma</u>: I_p is l.s.c. and convex.

<u>Proof</u>: Suppose that $I_p(\mu) > c$. Choose $A \in \mathring{\mathbf{C}}$ so that $\mu \in A$ and $\ell(A) < -c$. Then $-I_p(\nu) \leq \ell(A) < -c$ for all $\nu \in A$. Hence $\{\mu : I_p(\mu) > c\}$ is open. Thus I_p is l.s.c.

Next suppose that $\nu_1, \nu_2 \in \mathcal{M}_1(E)$ and that $\nu = \frac{\nu_1+\nu_2}{2}$. We want to show that $I(\nu) \leq 1/2(I(\nu_1) + I(\nu_2))$. Without loss in generality, we assume that $I(\nu_1) \vee I(\nu_2) < \infty$. Given $A \in \mathring{\mathbf{C}}$ with $\nu \in A$, choose $A_1, A_2 \in \mathring{\mathbf{C}}$ such

that $\nu_i \in A_i$ and $\operatorname{dist}(\frac{\bar{A}_1+\bar{A}_2}{2}, A^c) > 0$. Since $I(\nu_1) \vee I(\nu_2) < \infty$, $\ell(A_i) = \lim_{t\to\infty} \frac{1}{t} \log \mathcal{P}_t(A_i) > -\infty$ for $i = 1$ and 2. Hence $\mathcal{P}_{2t}(\frac{A_1+A_2}{2}) \geq \mathcal{P}_t(A_1)\mathcal{P}_t(A_2) > 0$ for large t's ; and so:

$$\begin{aligned}\ell(A) &= \lim_{t\to\infty} \frac{1}{2t} \log \mathcal{P}_{2t}(A) \\ &\geq 1/2 \lim_{t\to\infty} (\frac{1}{t} \log \mathcal{P}_t(A_1) + \frac{1}{t} \log \mathcal{P}_t(A_2)) \\ &= 1/2(\ell(A_1) + \ell(A_2)) \geq -1/2(I_p(\nu_1) + I_p(\nu_2)) .\end{aligned}$$

Since this is true for every $A \in \overset{\circ}{C}$ with $\nu \in A$, we conclude that $I_p(\nu) \leq 1/2(I_p(\nu_1) + I_p(\nu_2))$. □

(6.13) Lemma: Let $\mathcal{K} = \{K(\delta) : 0 < \delta < 1\}$ be a family of compact subsets of E and denote by $\mathcal{M}_1(\mathcal{K})$ the set of $Q \in \mathcal{M}_1(\mathcal{M}_1(E))$ satisfying $Q(\{\nu : \nu(K(\delta)^c) \geq \delta\}) \leq \delta$ for all $0 < \delta < 1$. Then, for each $L > 0$ there is a compact convex $C(L)$ in $\mathcal{M}_1(E)$ such that $Q^n(\frac{1}{n}\sum_1^n \nu_k \notin C(L)) \leq 2e^{-nL}$ for all $n \geq 1$ and $Q \in \mathcal{M}_1(\mathcal{K})$.

Proof: Let $0 < \varepsilon \leq 1/2$ and $0 < \delta < \varepsilon$ be given numbers which satisfy $\varepsilon \log \varepsilon/\delta \geq \log 2$. Define $G(\varepsilon,\delta) = \{\nu \in \mathcal{M}_1(E) : \nu(K(\delta)^c) \leq \varepsilon\}$. Since $\nu(K(\delta)^c) \leq \delta + \chi_{[\delta,1]}(\nu(K(\delta)^c))$,

$$\begin{aligned}&Q^n(\frac{1}{n}\sum_1^n \nu_k \notin G(2\varepsilon,\delta)) \\ &= Q^n(\frac{1}{n}\sum_1^n \nu_k(K(\delta)^c) > 2\varepsilon) \\ &\leq Q^n(\frac{1}{n}\sum_1^n \chi_{[\delta,1]}(\nu_k(K(\delta)^c)) > \varepsilon) .\end{aligned}$$

Hence, by the same argument as was used to prove (3.33), we conclude that for $Q \in \mathcal{M}_1(K)$ and $n \geq 1$:

$$Q^n\left(\frac{1}{n}\sum_1^n \nu_k \notin G(2\varepsilon,\delta)\right) \leq (\delta/\varepsilon)^{n\varepsilon/2} \quad .$$

Now let $L > \log 2$ be given. Set $\varepsilon_\ell = \frac{1}{2\ell}$ for $\ell \geq 1$ and define $\delta_\ell = \varepsilon_\ell e^{-2\ell L/\varepsilon_\ell}$. Then

$$Q^n\left(\frac{1}{n}\sum_1^n \nu_k \notin G(2\varepsilon_\ell,\delta_\ell)\right) \leq e^{-n\ell L}$$

for all $Q \in \mathcal{M}_1(K)$, $n \geq 1$, and $\ell \geq 1$. Hence, if $C(L) = \bigcap_{\ell=1}^{\infty} G(2\varepsilon_\ell,\delta_\ell)$, then

$$Q^n\left(\frac{1}{n}\sum_1^n \nu_k \notin C(L)\right) \leq 2e^{-nL}$$

for all $Q \in \mathcal{M}_1(K)$ and $n \geq 1$. Clearly $C(L)$ is compact and convex. □

(6.14) Lemma: Given $s \geq 1$ and $\alpha \geq 2$, define $J_{s,\alpha} : \Omega \to \mathcal{M}_1(E)$ by

$$J_{s,\alpha}(\Gamma) = \frac{2}{\alpha}\int_s^{s+\alpha/2} \chi_\Gamma(X(t))dt \quad , \quad \Gamma \in \mathcal{B}_E \quad .$$

Then for each $L > 0$ there is a compact convex $C(L)$ in $\mathcal{M}_1(E)$ such that

$$P_x\left(\frac{1}{n}\sum_{m=0}^{n-1} J_{s,\alpha} \circ \theta_{m\alpha} \notin C(L)\right) \leq 2e^{-nL}$$

for all $s \geq 1$, $\alpha \geq 2$, $n \geq 1$, and $x \in E$.

Proof: First observe that for any $t \geq 1$ and all $x_1, x_2 \in E$, $P(t,x_1,\cdot) \leq MP(1,x_2,\cdot)$. Indeed, since this is true when $t = 1$, if $t > 1$ we have:

$$P(t,x_1,\Gamma) = \int P(1,y,\Gamma)P(t-1,x_1,dy)$$
$$\leq MP(1,x_2,\Gamma) \quad .$$

Next, let $x_0 \in E$ be a fixed reference point and define $Q_{s,\alpha} = P_{x_0} \circ (J_{s,\alpha})^{-1}$. Then, since $s \geq 1$ and $\alpha \geq 2$, it follows from the preceding that for any $n \geq 1$ and any bounded measurable $F : (\mathfrak{M}_1(E))^n \to [0,\infty)$:

$$\sup_{x\in E} E^{P_x}[F(J_{s,\alpha},\ldots,J_{s,\alpha} \circ \theta_{(n-1)\alpha})]$$
$$\leq M^n E^{Q^n_{s,\alpha}}[F(\nu_1,\ldots\ldots,\nu_n)] \quad .$$

Thus, in view of Lemma (6.13) , we will be done as soon as we show that there is a family $\mathcal{K} = \{K(\delta) : 0 < \delta < 1\}$ of compact subsets of E such that $\{Q_{s,\alpha} : s \geq 1 \text{ and } \alpha \geq 2\} \subseteq \mathfrak{M}_1(\mathcal{K})$. But for given $s \geq 1$ and $\alpha \geq 2$, note that

$$E^{Q_{s,\alpha}}[\nu(\Gamma)] = \frac{2}{\alpha} E^{P_{x_0}}[\int_s^{s+\alpha/2} \chi_\Gamma(X(t))dt]$$
$$= \frac{2}{\alpha} \int_s^{s+\alpha/2} P(t,x_0,\Gamma)dt \leq MP(1,x_0,\Gamma) \quad .$$

Thus, if we take $K(\delta) \subset\subset E$ so that $P(1,x_0,K(\delta)^c) < \delta^2/M$, then for all $s \geq 1$ and $\alpha \geq 2$:

$$Q_{s,\alpha}(\nu(K(\delta)^c) \geq \delta) \leq 1/\delta \; E^{Q_{s,\alpha}}[\nu(K(\delta)^c)] < \delta \quad . \qquad \square$$

(6.15) <u>Lemma</u>: For each $L > 0$ there is a compact $C(L)$ in $\mathfrak{M}_1(E)$ such that

$$\overline{\lim_{T\to\infty}} \frac{1}{T} \log(\sup_{x\in E} P_x(L_T^1 \notin C(L))) \leq -L \quad .$$

Proof: Given $T \geq 4$, set $\alpha = T/[T/2]$, $n = [T/2]$, and note that

$$L_T^1 = 1/2(\frac{1}{n}\sum_{m=0}^{n-1}(J_{1,\alpha} + J_{1+\alpha/2,\alpha}) \circ \theta_{m\alpha}) \quad .$$

Hence, if $C(L)$, $L > 0$, is as in Lemma (6.14), then, since $C(4L)$ is convex:

$$\begin{aligned} &P_x(L_T^1 \notin C(4L)) \\ &\leq P_x(\frac{1}{n}\sum_{m=0}^{n-1} J_{1,\alpha} \circ \theta_{m\alpha} \notin C(4L)) \\ &\quad + P_x(\frac{1}{n}\sum_{m=0}^{n-1} J_{1+\alpha/2,\alpha} \circ \theta_{m\alpha} \notin C(4L)) \\ &\leq 4\exp(-4nL) \leq 4e^{-TL} \quad . \end{aligned}$$

□

Proof of (6.9) with $I = I_P$:

Introduce a metric ρ on $\mathcal{M}_1(E)$ having the property that $\rho(\alpha\nu_1 + (1-\alpha)\nu_2,\mu) \leq \alpha\rho(\nu_1,\mu) + (1-\alpha)\rho(\nu_2,\mu)$. (For example let $\{f_n\}_1^\infty \subseteq C_b(E)$ be a determining set of functions for convergence in $C_b(E)$, normalize the f_n's so that $\|f_n\|_{C_b(E)} = 1$, and set $\rho(\nu,\mu) = \sum_1^\infty \frac{1}{2^n}\left|\int f_n d\nu - \int f_n d\mu\right|$.) Then, for any $\nu \in \mathcal{M}_1(E)$ and $\varepsilon > 0$, the ρ-ball $B(\nu,\varepsilon) = \{\mu : \rho(\mu,\nu) < \varepsilon\}$ is convex.

We first prove the lower bound for open sets G in $\mathcal{M}_1(E)$. To this end, let $\nu \in G$. If $I_P(\nu) = \infty$, there is nothing to do. If $I_P(\nu) < \infty$, then for all $\varepsilon > 0$, $\ell(B(\nu,\varepsilon)) = \sup_{t>0} \frac{1}{t} \log\mathcal{P}_t(B(\nu,\varepsilon)) = \lim_{t\to\infty} \frac{1}{t} \log\mathcal{P}_t(B(\nu,\varepsilon))$. In particular, if $\varepsilon > 0$ is chosen so that $B(\nu,\varepsilon) \subseteq G$, then

$$-I_P(\nu) \le \ell(B(\nu,\varepsilon)) = \lim_{t\to\infty} \frac{1}{t} \log \mathcal{P}_t(B(\nu,\varepsilon)) \le \varliminf_{t\to\infty} \frac{1}{t} \log(\inf_{x\in E} Q_{t,x}(G)) \quad .$$

We next show that $\{I_P \le L\}$ is compact. To this end, choose $C(L)$ as in Lemma (6.15), and note that $\sup_{x\in E} Q^1_{t,x}(C(L)^c) \ge \inf_{x\in E} Q^1_{t,x}(C(L)^c) \ge \inf_{x\in E} Q_{t,x}(C(L)^c)$. Hence:

$$\begin{aligned} -L &\ge \varlimsup_{t\to\infty} \frac{1}{t} \log(\sup_{x\in E} Q^1_{t,x}(C(L)^c)) \\ &\ge \varliminf_{t\to\infty} \frac{1}{t} \log(\inf_{x\in E} Q_{t,x}(C(L)^c)) \\ &\ge -\inf_{\nu\in C(L)} I_P(\nu) \quad . \end{aligned}$$

From this it follows easily that $\{I_P \le L\} \subseteq C(L+1)$; and therefore, since $\{I_P \le L\}$ is closed, it is also compact.

We now prove the upper bound for compact sets K. Given $c > -\inf_{\nu\in K} I_P(\nu)$, choose a finite set $\nu_1,\dots,\nu_N \in K$ and positive numbers $\varepsilon_1,\dots,\varepsilon_N > 0$ so that $K \subseteq \bigcup_1^N B(\nu_i,\varepsilon_i)$ and $\ell(B(\nu_i,4\varepsilon_i)) < c$, $1 \le i \le N$. Clearly:

$$Q_{t,x}(K) \le N \max_{1\le i\le N} Q_{t,x}(B(\nu_i,\varepsilon_i)) \quad ,$$

and there is a $T > 0$ such that

$$\begin{aligned} Q_{t,x}(B(\nu_i,\varepsilon_i)) &\le Q^1_{t,x}(B(\nu_i,2\varepsilon_i)) \\ &\le M \inf_{y\in E} Q^1_{t,y}(B(\nu_i,2\varepsilon_i)) \\ &\le M\mathcal{P}_t(B(\nu_i,3\varepsilon_i)) \end{aligned}$$

for all $t \ge T$, $x \in E$, and $1 \le i \le N$. If $\max_{1\le i\le N} \sup_{t>0} \mathcal{P}_t(B(\nu_i,3\varepsilon_i)) = 0$, then

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(K)) = -\infty < c \quad .$$

If $\max_{1\le i\le N} \sup_{t>0} \mathcal{P}_t(B(\nu_i,\varepsilon_i)) > 0$, set $J = \{1 \le i \le N : \sup_{t>0} \mathcal{P}_t(B(\nu_i,\varepsilon_i)) > 0\}$. Then

$$\frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(K)) \le \max_{i\in J} \frac{1}{t} \log(MN\mathcal{P}_t(B(\nu_i,4\varepsilon_i)))$$

for $t \ge T$, and

$$c > \mathcal{L}(B(\nu_i,4\varepsilon_i)) = \lim_{t\to\infty} \frac{1}{t} \log(\mathcal{P}_t(B(\nu_i,4\varepsilon_i)))$$

for each $i \in J$. Hence, once again:

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(K)) < c \quad .$$

Observe that, since $\sup_{x\in E} Q^1_{t,x}(K) \le \sup_{x\in E} Q_{t,x}(K)$, we now also know that

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q^1_{t,x}(K)) \le -\inf_{\nu\in K} I_P(\nu)$$

for $K \subset\subset \mathcal{M}_1(\mathcal{E})$. Now suppose that F is closed and choose $C(L)$, $L > 0$, as in Lemma (6.15) . Then:

$$Q^1_{t,x}(F) \le Q^1_{t,x}(F \cap C(L)) + Q^1_{t,x}(C(L)^c) \quad .$$

By the preceding plus Lemma (6.15) , for any $\varepsilon > 0$ we have

$$Q^1_{t,x}(F \cap C(L)) + Q^1_{t,x}(C(L)^c)$$

$$\le 2\exp(-t((\inf_{\nu\in F} I_P(\nu) \wedge L) - \varepsilon))$$

for sufficiently large t and all $x \in E$. From this it is clear that

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q^1_{t,x}(F)) \leq -\inf_{\nu\in F} I_P(\nu) \quad .$$

Finally, let F be closed. Given $\varepsilon > 0$, we have:

$$\sup_{x\in E} Q_{t,x}(F) \leq \sup_{x\in E} Q^1_{t,x}(\overline{F^{(\varepsilon)}})$$

for sufficiently large t's. Hence, by the preceding,

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(F)) \leq -\inf_{\nu\in\overline{F^{(\varepsilon)}}} I_P(\nu)$$

for all $\varepsilon > 0$. But, from what we already know about I_P, we have $\inf_{\nu\in F} I_P(\nu) = \lim_{\varepsilon\downarrow 0} \inf_{\nu\in F^{(\varepsilon)}} I_P(\nu)$ (cf. remark (2.4)). □

7. Identification of the Rate Function:

We want to carry out the program outlined in section 5), now that we know some conditions which guarantee the existence of a rate function. Throughout this section we will be working with one or the other of the following situations.

(D) $\{P_x : x \in E\}$ is a discrete time, Feller continuous, time-homogeneous Markov family on $\Omega = E^{\mathcal{N}}$ with transition function $\pi(x,\cdot)$. Define $\pi : C_b(E) \to C_b(E)$ by $[\pi\phi](x) = \int\phi(y)\pi(x,dy)$; and, for each $V \in C_b(E)$, define $\pi_V\phi(x) = e^{V(x)}[\pi\phi](x)$.

(C) $\{P_x : x \in E\}$ is a continuous time, Feller continuous, time-homogeneous Markov family on $\Omega = D([0,\infty);E)$ with transition function $P(t,x,\cdot)$. Define $P_t : C_b(E) \to C_b(E)$ by $[P_t\phi](x) = \int\phi(y)P(t,x,dy)$; and, for each $V \in C_b(E)$, let P^V_t denote the unique semi-group on $C_b(E)$ satisfying

$$P^V_t = P_t + \int_0^t P^V_{t-s}(VP_s)ds \quad , \quad t \geq 0 \quad . \tag{7.1}$$

Let A with domain $D(A)$ denote the weak generator of $\{P_t : t > 0\}$.

(7.2) Lemma: In case (D) :

$$(7.3) \qquad \pi_V^n\phi(x) = E^{P_x}[\exp(\sum_0^{n-1} V(X(m)))\phi(X(n))] \quad , \quad \phi \in C_b(E) \quad ,$$

for each $V \in C_b(E)$. In case (C) :

$$(7.4) \qquad P_t^V\phi(x) = E^{P_x}[\exp(\int_0^t V(X(s))ds)\phi(X(t))] \quad , \quad \phi \in C_b(E) \quad ,$$

for each $V \in C_b(E)$. Moreover, in case (C) , if A_V denotes the weak generator of $\{P_t^V : t > 0\}$, then A_V has domain $D(A)$ and $A_V\phi = A\phi + V\phi$ for $\phi \in D(A)$.

Proof: (7.3) and (7.4) were proved in Lemmas (5.7) and (5.15) , respectively. To see that $D(A)$ is contained in the domain of A_V and that $A_V\phi = A\phi + V\phi$, one simply uses (7.1) to compute $\lim_{t\downarrow 0}(P_t^V\phi-\phi)/t$ for $\phi \in D(A)$. To prove that the domain of A_V is contained in $D(A)$, one uses (7.1) to compute $\lim_{t\downarrow 0}(P_t\phi-\phi)/t$ for ϕ in the domain of A_V . ☐

We next define $\lambda(V)$, $V \in C_b(E)$, by:

$$(7.5) \qquad \lambda(V) = \begin{cases} \lim_{n\to\infty} \frac{1}{n} \log(\|\pi_V^n\|_{op}) & \text{in case (D)} , \\ \lim_{t\to\infty} \frac{1}{t} \log(\|P_t^V\|_{op}) & \text{in case (C)} , \end{cases}$$

Note that the indicated limits exist in each case, since $n \to \log(\|\pi_V^n\|_{op})$ and $t \to \log(\|P_t^V\|_{op})$ are both sub-additive. Also:

$$(7.6) \qquad \|\pi_V^n\|_{op} = \|E^{P_\bullet}[\exp(\sum_0^{n-1} V(X(m)))]\|_{C_b(E)}$$

and

$$\|P_t^V\|_{op} = \|E^{P_\bullet}[\exp(\int_0^t V(X(s))ds)]\|_{C_b(E)} \tag{7.7}$$

(7.8) <u>Lemma</u>: The map $\lambda : C_b(E) \to R^1$ is continuous and convex.

<u>Proof</u>: First using (7.6) and (7.7), note that $|\lambda(V)| \leq \|V\|_{C_b(E)}$. Next, from (7.6) and Hölder's inequality, $\|\pi^n_{\theta V_1+(1-\theta)V_2}\|_{op} \leq \|\pi^n_{V_1}\|^{\theta}_{op}\|\pi^n_V\|^{(1-\theta)}_{op}$; for $\theta \in (0,1)$ and $V_1, V_2 \in C_b(E)$. Similarly, $\|P_t^{\theta V_1+(1-\theta)V_2}\|_{op} \leq \|P_t^{V_1}\|^{\theta}_{op}\|P_t^{V_2}\|^{(1-\theta)}_{op}$. Thus λ is convex. Finally, given $V_1, V_2 \in C_b(E)$, set $V = (V_2-V_1)/\varepsilon$ where $\varepsilon = \|V_2-V_1\|_{C_b(E)}$. Assume that $0 < \varepsilon < 1$. Then, $\lambda(V_2) = \lambda((1-\varepsilon)V_1 + \varepsilon(V+V_1)) \leq (1-\varepsilon)\lambda(V_1) + \varepsilon\lambda(V+V_1)$; and so $\lambda(V_2) - \lambda(V_1) \leq \varepsilon(|\lambda(V_1)| + \lambda(V+V_1)) \leq (2\|V_1\|_{C_b(E)}+1)\|V_2-V_1\|_{C_b(E)}$. Reversing the roles of V_1 and V_2, we get

$$|\lambda(V_2)-\lambda(V_1)| \leq (2\|V_1\|_{C_b(E)} \vee \|V_2\|_{C_b(E)}+1)\|V_2-V_1\|_{C_b(E)} . \quad \square$$

Now let $\mathcal{M}(E)$ be the space of all finite signed Borel measures on E and define $\lambda^* : \mathcal{M}(E) \to [0,\infty) \cup \{\infty\}$ by

$$\lambda^*(\nu) = \sup_{V \in C_b(E)} (\int V d\nu - \lambda(V)) . \tag{7.9}$$

(7.10) <u>Lemma</u>: The map λ^* is a l.s.c. convex function on $\mathcal{M}(E)$ (with the weak topology). Moreover, $\lambda^*(\nu) = \infty$ for $\nu \notin \mathcal{M}_1(E)$.

<u>Proof</u>: As the supremum of continuous linear functions, λ^* is l.s.c. and convex.

To prove that $\lambda^*(\nu) = \infty$ for $\nu \notin \mathcal{M}_1(E)$, first suppose that

$\nu = \nu_+ - \nu_-$ where $\nu_+, \nu_- \in \mathcal{M}(E)^+$ and $\nu_- \neq 0$. Then there is a $V \in C_b(E)^+$ such that $\int V d\nu \leq -1$. Since $\lambda(-\alpha V) \leq 0$ for all $\alpha > 0$, $\lambda^*(\nu) \geq \int(-\alpha V)d\nu - \lambda(-\alpha V) \geq \alpha$ for all $\alpha > 0$. Hence $\lambda^*(\nu) = \infty$ in this case. Next, suppose that $\nu \in \mathcal{M}(E)^+$ and that $\nu(E) \neq 1$. Then, since $\lambda(\alpha 1) = \alpha$ for all $\alpha \in R^1$, $\lambda^*(\nu) \geq \alpha\nu(E) - \alpha = (\nu(E) - 1)\alpha$ for all $\alpha \in R^1$. Thus, $\lambda^*(\nu) = \infty$ in this case also.

□

We define $J : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ by:

$$(7.11) \qquad J(\nu) = \begin{cases} -\inf\{\int \log \frac{\pi u}{u} d\nu : u \in \mathcal{U}\} & \text{in case } (\underline{D}) \\ -\inf\{\int \frac{Au}{u} d\nu : u \in \mathcal{U} \cap D(A)\} & \text{in case } (\underline{C}) \end{cases} ,$$

where $\mathcal{U} = \{u \in C_b(E) : (\exists \varepsilon > 0)\ u \geq \varepsilon\}$. We extend J to $\mathcal{M}(E)$ by setting $J(\nu) = \infty$ for $\nu \notin \mathcal{M}_1(E)$. Since J is clearly l.s.c. and convex on $\mathcal{M}_1(E)$ and because $\mathcal{M}_1(E)$ is closed and convex in $\mathcal{M}(E)$, J is l.s.c. and convex on $\mathcal{M}(E)$. Our goal is to prove that $\lambda^* = J$. Half of this equality is easy. Namely, given $u \in \mathcal{U}$ $(u \in \mathcal{U} \cap D(A))$, define $V_u = -\log\frac{\pi u}{u}$ $(V_u = -\frac{Au}{u})$. Then $\pi_{V_u} u = u$ $(A_{V_u} u = 0)$. Hence $\pi^n_{V_u} u = u$ $(P^{V_u}_t u = u)$; and, therefore, since $u \geq \varepsilon$ for some $\varepsilon > 0$, it is easy to see that $\lambda(V_u) = 0$. Hence, if $\nu \in \mathcal{M}_1(E)$, then

$$\lambda^*(\nu) \geq \sup_{u \in \mathcal{U}} \int V_u d\nu = \sup_{u \in \mathcal{U}}(-\int \log \frac{\pi u}{u} d\nu) = J(\nu)$$

$$(\lambda^*(\nu) \geq \sup_{u \in \mathcal{U} \cap D(A)} \int V_u d\nu = \sup_{u \in \mathcal{U} \cap D(A)} (-\int \frac{Au}{u} d\nu) = J(\nu))$$

That is,

$$(7.12) \qquad \lambda^* \geq J \quad .$$

The proof of the opposite inequality will require us to take an excursion into

some abstract convex analysis.

Lemma (7.13) and Theorem (7.15), which follow, refer to the following set-up. The space B is a locally convex, Hausdorff topological vector space over the reals; and B^* is the dual of B. Given $f : B \to R^1 \cup \{\infty\}$, the <u>epi-graph</u> of f is the set <u>epi(f)</u> $= \{(b,\mu) \in B \times R^1 : \mu \geq f(b)\}$. The <u>dual epi-graph</u> of f is the set <u>$\text{epi}^*(f)$</u> $= \{(b^*,\mu^*) \in B^* \times R^1 : f(b) \geq b^*(b) - \mu^*$ for all $b \in B\}$. The next lemma is an immediate consequence of the preceding definitions.

(7.13) <u>Lemma</u>: If $f : B \to R^1 \cup \{\infty\}$, then

<u>a</u>) epi(f) is closed if and only if f is l.s.c.

<u>b</u>) epi(f) is convex if and only if f is convex.

Moreover, $\text{epi}^*(f)$ is a closed convex subset of $B^* \times R^1$. Finally, assuming that $f \not\equiv \infty$, define $f^* : B^* \to R^1 \cup \{\infty\}$ by

$$f^*(b^*) = \sup_{b \in B}(b^*(b) - f(b)) \quad . \tag{7.14}$$

Then f^* is l.s.c., convex, and $\text{epi}^*(f) = \text{epi}(f^*)$.

The function f^* is called the <u>conjugate convex function</u> of f.

(7.15) <u>Theorem</u>: If $f : B \to R^1 \cup \{\infty\}$ is a l.s.c, convex function which is not identically equal to $+\infty$, then $f(b) = \sup_{b^* \in B^*} (b^*(b) - f^*(b^*))$ for all $b \in B$.

<u>Proof</u>: We first prove that

$$f(b) = \sup\{b^*(b) - \mu^* : (b^*,\mu^*) \in \text{epi}^*(f)\} \quad . \tag{7.16}$$

To this end, note that since epi(f) is closed and convex,

$\mathrm{epi}(f) = \bigcap\{H : H \in \mathscr{H}\}$ where $\mathscr{H}$ is the set of all closed affine half spaces in $B \times R^1$, containing $\mathrm{epi}(f)$. (This is a corollary of the geometric form of the Hahn-Banach Theorem and can be found in most modern treatments of abstract functional analysis. See, for example, [Fnal. Anal., L. Schwartz]. Next, if H is a closed affine half space in $B \times R^1$, then H is determined by a triple $(b^*, \rho^*, \mu^*) \in B^* \times R^1 \times R^1$ such that $(b^*, \rho^*) \neq \{0,0\}$ and $H = \{(b,\xi) \in B \times R^1 : b^*(b) - \rho^* \xi \leq \mu^*\}$. Moreover, the triple (b^*, ρ^*, μ^*) is determined by H up to a positive multiplication factor. In particular, if $H \in \mathscr{H}$ and H is determined by (b^*, ρ^*, μ^*), then, choosing $b_0 \in B$ so that $f(b_0) < \infty$, we see that $b^*(b_0) - \rho^* \xi \leq \mu^*$ for all $\xi \geq f(b_0)$. Thus $\rho^* \geq 0$. We partition $\mathscr{H}$ into $\mathscr{H}^+$ and $\mathscr{H}^0$ according to whether $\rho^* > 0$ or $\rho^* = 0$. For each $H \in \mathscr{H}^+$, there is a unique $(b_H^*, \mu_H^*) \in B^* \times R^1$ such that $H = \{(b,\xi) : b_H^*(b) - \xi \leq \mu_H^*\}$.

Note that $\mathscr{H}^+ \neq \emptyset$. Indeed, if $\mathscr{H}^+ = \emptyset$, then we would have $\mathrm{epi}(f) = \bigcap\{H : H \in \mathscr{H}^0\}$. Since each $H \in \mathscr{H}^0$ has the property that $(b,\xi) \in H$ for all $\xi \in R^1$ whenever $(b,\xi) \in H$ for some $\xi \in R^1$, $\mathrm{epi}(f)$ would have the same property. Since $\mathrm{epi}(f) \neq \emptyset$, and $f > -\infty$, this cannot be. Hence, $\mathscr{H}^+ \neq \emptyset$.

We next need to show that

$$\mathrm{epi}(f) = \bigcap\{H : H \in \mathscr{H}^+\} . \tag{7.17}$$

Clearly (7.17) will follow once we show that for any $H_0 \in \mathscr{H}^0$ and $(b_0, \xi_0) \notin H_0$, there is an $H \in \mathscr{H}^+$ such that $(b_0, \xi_0) \notin H$. To this end, choose $b_0^* \in B^* \setminus \{0\}$ and μ_0^* so that $H_0 = \{b : b_0^*(b) \leq \mu_0^*\} \times R^1$. Next, choose $H_1 \in \mathscr{H}^+$, and for $\lambda > 0$ define

$$H_\lambda = \{(b,\xi) : (\lambda b_0^* + b_{H_1}^*)(b) - \xi \leq \lambda\mu_0^* + \mu_{H_1}\} \quad .$$

Then $H_\lambda \in \mathcal{H}^+$ for each $\lambda > 0$. Indeed, if $(b,\xi) \in \text{epi}(f)$, then $b_0^*(b) \leq \mu_0^*$ and $b_{H_1}^*(b) - \xi \leq \mu_{H_1}^*$. Next, since $(b_0,\xi_0) \notin H_0$, $b_0^*(b_0) > \mu_0^*$ and so $b_{H_\lambda}^*(b_0) - \mu_{H_\lambda}^*$ can be made larger than ξ_0 simply by taking λ sufficiently large. Thus, for large enough λ , $(b_0,\xi_0) \notin H_\lambda$.

We can now prove (7.16) . Certainly, $f(b) \geq \sup\{b^*(b) - \mu^* : (b^*,\mu^*) \in \text{epi}^*(f)\}$. On the other hand, if $\xi < f(b)$, we can find an $H \in \mathcal{H}^+$ so that $(b,\xi) \notin H$. Hence, $b_H^*(b) - \xi > \mu_H^*$. This proves that $f(b) \leq \sup\{b_H^*(b) - \mu_H^* : H \in \mathcal{H}^+\}$. But $H \in \mathcal{H}^+$ implies that $(b,f(b)) \in H$ for all b satisfying $f(b) < \infty$; and so $f(b) \geq b_H^*(b) - \mu_H^*$ for all b . In other words, $(b_H^*,\mu_H^*) \in \text{epi}^*(f)$ for all $H \in \mathcal{H}^+$. (7.16) is therefore proved.

Finally, since $f^*(b^*) = \inf\{\mu^* : (b^*,\mu^*) \in \text{epi}^*(f)\}$ and $(b^*,f^*(b^*)) \in \text{epi}^*(f)$ so long as $f^*(b^*) < \infty$, our theorem is an immediate consequence of (7.16) .

□

(7.18) <u>Theorem</u>: $\lambda^* = J$.

<u>Proof</u>: We have already seen that $\lambda^* \geq J$ (cf. (7.12)). To prove the opposite inequality, define $J^*(V) = \sup_{\nu \in \mathcal{M}(E)} (\int V d\nu - J(\nu))$ for $V \in C_b(E)$. Since $C_b(E) = \mathcal{M}(E)^*$ (cf. the proof of Lemma (3.37)),

$$J(\nu) = \sup_{V \in C_b(E)} (\int V d\nu - J^*(V)) \quad .$$

This follows from Theorem (7.15) if $J \not\equiv \infty$. If $J \equiv \infty$, it is trivial, since then $J^* \equiv -\infty$. Thus, if we show that $J^* \leq \lambda$, then we will know that

$$\lambda^*(\nu) = \sup_V(\int V d\nu - \lambda(V)) \le$$

$$\sup_V(\int V d\nu - J^*(V)) = J(\nu) \quad .$$

We will now show that $J^* \le \lambda$.

In case $(\underline{D})$, we proceed as follows. Given $\lambda > \lambda(V)$, set $u_\lambda = \sum_0^\infty e^{-\lambda n}[\pi_V^n 1]$. Then, since $\|\pi_V^n 1\|_{C_b(E)} \le e^{n(\lambda-\varepsilon)}$ for some $\varepsilon > 0$ and all large n's and $u_\lambda \ge 1$, we see that $u_\lambda \in \mathcal{U}$. Moreover, $\pi_V u_\lambda = e^\lambda(u_\lambda - 1)$; and so $\log \frac{\pi_V u_\lambda}{u_\lambda} \le \lambda$. In particular,

$$\sup_{\nu \in \mathcal{M}_1(E)} \int \log \frac{\pi_V u_\lambda}{u_\lambda} d\nu \le \lambda \quad .$$

We have now shown that

$$\lambda(V) \ge \sup_{\nu\in\mathcal{M}_1(E)} \inf_{u\in U} \int \log \frac{\pi_V u}{u} d\nu = \sup_{\nu\in\mathcal{M}_1(E)} (\int V d\nu + \inf_{u\in\mathcal{U}} \int \log \frac{\pi u}{u} d\nu)$$

$$= \sup_{\nu\in\mathcal{M}(E)} (\int V d\nu - J(\nu)) = J^*(V) \quad .$$

In case $(\underline{C})$, if $\lambda > \lambda(V)$ we set $u_\lambda = \int_0^\infty e^{-\lambda t}[P_t^V 1]dt$. Then, as in the preceding, $u_\lambda \in \mathcal{U}$. Moreover,

$$P_s^V u_\lambda = e^{\lambda s}\int_s^\infty e^{-\lambda t}[P_t^V 1]dt = e^{\lambda s}(u_\lambda - \int_0^s e^{-\lambda t}[P_t^V 1]dt)$$

and so $u_\lambda \in D(A)$ and $A_V u_\lambda = \lambda u_\lambda - 1$. Hence $\frac{A_V u_\lambda}{u_\lambda} \le \lambda$. The rest is just like the preceding. □

(7.19) <u>Corollary</u>: If E is compact, then

$$(7.20) \qquad \lambda(V) = \sup_{\nu \in \mathcal{M}_1(E)} \left(\int V d\nu - J(\nu) \right) .$$

Proof: Since E is compact, $C_b(E)^* = \mathcal{M}(E)$. Therefore, (7.20) follows from Theorem (7.15) and (7.18) . □

Define $\{Q_{n,x} : n \geq 1$ and $x \in E\}$ in case (D) and $\{Q_{t,x} : t > 0$ and $x \in E\}$ in case (C) as in section 6) . Let $I : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ be a rate function. We say that $\{Q_{n,x} : n \geq 1$ and $x \in E\}$ ($\{Q_{t,x} : t > 0$ and $x \in E\}$) satisfy the uniform large deviation principle with rate I if

$$\underline{\lim}_{n\to\infty} \frac{1}{n} \log(\inf_{x\in E} Q_{n,x}(G)) \geq -\inf_{\nu\in G} I(\nu)$$

$$\left(\underline{\lim}_{t\to\infty} \frac{1}{t} \log(\inf_{x\in E} Q_{t,x}(G)) \geq -\inf_{\nu\in G} I(\nu)\right)$$

for all open G in $\mathcal{M}_1(E)$ and

$$\overline{\lim}_{n\to\infty} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(F)) \leq -\inf_{\nu\in F} I(\nu)$$

$$\left(\overline{\lim}_{t\to\infty} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(F)) \leq -\inf_{\nu\in F} I(\nu)\right)$$

for all closed F in $\mathcal{M}_1(E)$.

(7.21) Corollary: Suppose that $\{Q_{n,x} : n \geq 1$ and $x \in E\}$ in case (D) or $\{Q_{t,x} : t > 0$ and $x \in E\}$ in case (C) satisfies the uniform large deviation principle with some convex rate I . Then $I = J$. Moreover, in case (D) , one has, for each $\Phi \in C_b(E)$, that

$$(7.22) \qquad \lim_{n\to\infty} \sup_{x\in E} \left| \frac{1}{n} \log E^{Q_{n,x}}[\exp(n\phi)] - \sup_{\nu}(\Phi(\nu) - J(\nu)) \right| = 0 ;$$

and in case $(\underline{C})$:

(7.23) $$\lim_{t\to\infty} \sup_{x\in E} \left| \frac{1}{t} \log E^{Q_{t,x}}[\exp(t\phi)] - \sup_{\nu}(\Phi(\nu) - J(\nu)) \right| = 0 .$$

Proof: Repeating the argument used to prove Theorem (2.6), one sees that (7.22) and (7.23) follow easily with J replaced by I. Thus we need only show that $I = J$. But if $\Phi(\nu) = \int V d\nu$, then (7.22) and (7.23), with J replaced by I, says that $\lambda(V) = I^*(V)$. Hence, $I = \lambda^* = J$.

Another useful consequence of Theorem (7.18) is the following remark.

(7.24) Corollary: In case $(\underline{C})$, for each $h > 0$ define $J_h(\nu) = -\inf_{u\in\mathcal{U}} \int \log\left(\frac{P_h u}{u}\right) d\nu$, $\nu \in \mathcal{M}_1(E)$. Then $J_h \leq hJ$.

Proof: Define $\lambda_h(V) = \lim_{n\to\infty} \frac{1}{n} \log\|((P_h)^V)^n\|_{op}$. Then $J_h = \lambda_h^*$. Suppose that we could show that $\lambda_h(V) \geq h\lambda(V/h)$, $V \in C_b(E)$. Then we would have:

$$J_h(\nu) \leq \sup_V\left(\int V d\nu - h\lambda(V/h)\right) = h\lambda^*(\nu) = hJ(\nu) .$$

To show that $\lambda_h(V) \geq h\lambda(V/h)$, note that for any $n \geq 1$ and $x \in E$:

$$E^{P_x}\left[\exp\left(\int_0^{nh} V(X(t))dt\right)\right]$$

$$= E^{P_x}\left[\exp\left(\sum_{k=0}^{n-1} \int_0^1 hV(X(kh+th))dt\right)\right]$$

$$\leq \int_0^1 E^{P_x}\left[\exp\left(\sum_{k=0}^{n-1} hV(X(kh+th))\right)\right]dt$$

$$= \int_0^1 E^{P_x}[E^{P_{x(th)}}[\exp(\sum_{k=0}^{n-1} hV(X(kh)))]]dt$$

$$= \int_0^1 [P_{th} \circ ((P_h)^{hV})^n 1](x)dt$$

$$\leq \|((P_h)^{hV})^n\|_{op} \quad ,$$

where we have used the fact that $(\xi_0,\ldots,\xi_{n-1}) \to \exp(\sum_0^{n-1} \xi_k)$ is convex in order to pass from the line two to line three. Hence

$$\|P^V_{nh}\|_{op} \leq \|((P_h)^{hV})^n\|_{op} \quad , \quad n \geq 1 \quad ,$$

and so

$$\lambda(V) = \lim_{n\to\infty} \frac{1}{nh} \log\|P^V_{nh}\|_{op} \leq \frac{1}{h} \lim_{n\to\infty} \frac{1}{n} \log\|((P_h)^{hV})^n\|_{op}$$

$$= \frac{1}{h} \lambda_h(hV) \quad .$$

Clearly, this is equivalent to $\lambda_h(V) \geq h\lambda(V/h)$. □

The main application of Corollary (7.24) is the next result.

(7.25) <u>Theorem</u>: There is a universal non-decreasing function $\rho : [0,\infty] \to [0,2]$ such that $\rho(0) = \lim_{\ell\downarrow 0} \rho(\ell) = 0$, $\lim_{\ell\to\infty} \rho(\ell) = 2$, and

$$\begin{cases} \|\nu\pi-\nu\|_{var} \leq \rho(J(\nu)) & \text{in case } (\underline{D}) \\ \|\nu P_h-\nu\|_{var} \leq \rho(hJ(\nu)) & \text{in case } (\underline{C}) \end{cases} \quad .$$

<u>Proof</u>: In view of (7.24) , we need only work in case $(\underline{D})$. To this end, first note that, for $\nu_1,\nu_2 \in \mathcal{M}_1(E)$, $\|\nu_1-\nu_2\|_{var} = 2\sup\{|\int v d\nu_1 - \int v d\nu_2| : v \in C_b(E) \text{ and } 0 \leq v \leq 1\}$. Next, suppose that $\ell = J(\nu) < \infty$. Given $v \in C_b(E)$ satisfying $0 \leq v \leq 1$ and $\varepsilon > 0$, we have:

$$\int \log(1 + \varepsilon\pi v)d\nu - \int \log(1 + \varepsilon v)d\nu \geq -\ell \quad .$$

Since $\alpha \in (0,\infty) \to \log \alpha$ is concave, this yields:

$$\log(1 + \varepsilon\int vd(\nu\pi)) - \int \log(1 + \varepsilon v)d\nu \geq -\ell \quad .$$

But $\log(1+\alpha) \leq \alpha$ for $\alpha \geq 0$ and $\log(1 + \varepsilon y) \geq y \log(1+\varepsilon)$ for all $\varepsilon \geq 0$ and $y \in [0,1]$. Thus

$$\varepsilon\int vd(\nu\pi) \geq -\ell + (\log(1+\varepsilon))\int vd\nu \quad .$$

In other words:

$$\begin{aligned} \varepsilon(\int vd(\nu\pi) - \int vd\nu) &\geq -\ell - (\varepsilon - \log(1+\varepsilon))\int vd\nu \\ &\geq -\ell - (\varepsilon - \log(1+\varepsilon)) \end{aligned}$$

since $\varepsilon - \log(1+\varepsilon) \geq 0$ and $\int vd\nu \leq 1$. Noting that the same inequality holds for $(1-v)$ in place of v , we have now shown that

$$\left|\int vd(\nu\pi) - \int vd\nu\right| \leq (\ell + (\varepsilon-\log(1+\varepsilon)))/\varepsilon \quad ;$$

and so

$$\|\nu\pi-\nu\|_{var} \leq 2\inf_{\varepsilon>0}(\ell + (\varepsilon-\log(1+\varepsilon)))/\varepsilon \quad .$$

Take $\rho(\ell) = 2\inf_{\varepsilon>0}(\ell + (\varepsilon-\log(1+\varepsilon)))/\varepsilon$. □

(7.26) Corollary: Let $\nu \in \mathfrak{M}_1(E)$. In either case (D) or case (C) , $J(\nu) = 0$ if and only if ν is invariant (i.e., $\nu\pi = \nu$ in (D) and $\nu P_t = \nu$, $t > 0$, in (C)) .

Proof: The "only if" is immediate from Theorem (7.25) . To prove the

"if", set $P_\nu = \int P_x \nu(dx)$. Given $V \in C_b(E)$, we have, in case (D) :

$$\log E^{P_\nu}[\exp(\sum_0^{n-1} V(X(m)))] \geq \sum_0^{n-1} E^{P_\nu}[V(X(m))] = n\int V d\nu$$

and in case (C) :

$$\log E^{P_\nu}[\exp(\int_0^t V(X(s))ds)] \geq \int_0^t E^{P_\nu}[V(X(s))]ds$$

$$= t\int V d\nu \quad .$$

(We have used here Jensen's inequality and the invariance of ν .) Hence, in case (D) :

$$\lambda(V) \geq \overline{\lim_{n\to\infty}} \frac{1}{n} \log E^{P_\nu}[\exp(\sum_0^{n-1} V(X(m)))] \geq \int V d\nu$$

and in case (C)

$$\lambda(V) \geq \overline{\lim_{t\to\infty}} \frac{1}{t} \log E^{P_\nu}[\exp(\int_0^t V(X(s))ds)] \geq \int V d\nu \quad .$$

In particular, $J(\nu) = \sup_V(\int V d\nu - \lambda(V)) \leq 0$.

□

For purposes of identifying the rate function arising in non-uniform large deviation principles, we introduce a slightly different characterization of J . Namely, given $V \in C_b(E)$, define

$$\tilde{\lambda}(V) = \begin{cases} \sup_{x\in E} \overline{\lim_{n\to\infty}} \frac{1}{n} \log([\pi_V^n 1](x)) & \text{in case (D)} \\ \sup_{x\in E} \overline{\lim_{t\to\infty}} \frac{1}{t} \log([P_t^V 1](x)) & \text{in case (C)} \end{cases} \quad .$$

(7.27) Theorem: The function $\tilde{\lambda} : C_b(E) \to R^1$ is continuous and convex and satisfies $\tilde{\lambda} \leq \lambda$. Moreover, $J = (\tilde{\lambda})^*$.

Proof: The convexity of $\tilde{\lambda}$ follows easily from the convexity of $V \to [\pi_V^n 1](x) (V \to [P_t^V 1](x))$ for each $n \geq 1$ $(t > 0)$ and $x \in E$. The inequality $\tilde{\lambda} \leq \lambda$ is trivial, and the continuity follows in the same way as it did for λ.

To prove that $J = \tilde{\lambda}^*$, first note that $J = \lambda^* \leq (\tilde{\lambda})^*$, since $\tilde{\lambda} \leq \lambda$. We now prove the opposite inequality by checking that $J^* \leq \tilde{\lambda}$ and therefore that $J = (J^*)^* \geq (\tilde{\lambda})^*$.

In case (D), let $V \in C_b(E)$ and $\lambda > \tilde{\lambda}(V)$ be given. For $n \geq 1$, define $u_{n,\lambda} = \sum_0^n e^{-\lambda m}[\pi_V^m 1]$. Clearly $u_{n,\lambda} \in C_b(E)$ and $u_{n,\lambda} \geq 1 + e^{-(\|V\|+\lambda)}$. Thus $u_{n,\lambda} \in \mathcal{U}$. Moreover, $\pi_V u_{n,\lambda} = e^{\lambda}(u_{n,\lambda} - 1) + e^{-\lambda n}[\pi_V^{n+1} 1]$. Hence

$$\frac{\pi_V u_{n,\lambda}}{u_{n,\lambda}} = e^{\lambda}\left(1 - \frac{1}{u_{n,\lambda}}\right) + \frac{e^{-\lambda n}[\pi_V^{n+1} 1]}{u_{n,\lambda}} .$$

Noting that $e^{-\lambda n}[\pi_V^{n+1} 1] \leq e^{\|V\|} e^{-\lambda n}[\pi_V^n 1] \leq e^{\|V\|} u_{n,\lambda}$, we see that $\sup\limits_{n \geq 1} \sup\limits_{x \in E} \log \frac{\pi_V u_{n,\lambda}(x)}{u_{n,\lambda}(x)} < \infty$. Next, note that

$$\frac{\pi_V u_{n,\lambda}}{u_{n,\lambda}} \leq e^{\lambda}(1 + e^{-(n+1)\lambda}[\pi^{n+1} 1]) .$$

Since $\lambda > \tilde{\lambda}(V)$, $e^{-\lambda m}[\pi_V^m 1](x) \to 0$, $x \in E$, as $m \to \infty$. We therefore have that

$$\overline{\lim_{n\to\infty}} \log([\pi_V u_{n,\lambda}](x)/u_{n,\lambda}(x)) \leq \lambda$$

for all $x \in E$. By Fatou's lemma, it follows that

$$\inf_{u \in \mathcal{U}} \int \log\left(\frac{\pi_V u}{u}\right) d\nu \leq \overline{\lim_{n\to\infty}} \int \log\left(\frac{\pi_V u_{n,\lambda}}{u_{n,\lambda}}\right) d\nu \leq \lambda$$

for all $\nu \in \mathcal{M}_1(E)$ and $\lambda > \tilde{\lambda}(V)$. Hence

$$J^*(V) = \sup_{\nu \in \mathcal{M}_1(E)} \inf_{u \in \mathcal{U}} \int \log\left(\frac{\pi_V u}{u}\right) d\nu \leq \tilde{\lambda}(V) \quad .$$

In case $(\underline{C})$, let $V \in C_b(E)$ and $\lambda > \tilde{\lambda}(V)$ be given. For $n \geq 1$, define $u_{n,\lambda} = \int_0^n e^{-\lambda t} P_t^V 1 \, dt$. Then $u_{n,\lambda} \in C_b(E) \cap D(A)$, $u_{n,\lambda} \geq e^{-(\|V\|+\lambda)}$, and

$$A_V u_{n,\lambda} = \lambda u_{n,\lambda} - 1 + e^{-\lambda n} P_n^V 1 \quad .$$

Hence, $u_{n,\lambda} \in \mathcal{U} \cap D(A)$ and

$$\frac{A_V u_{n,\lambda}}{u_{n,\lambda}} \leq \lambda + \frac{e^{-\lambda n} P_n^V 1}{u_{n,\lambda}}$$

Since $\lambda > \tilde{\lambda}(V)$, and therefore $e^{-\lambda n} P_n^V 1 \to 0$ pointwise, we will be in a position to proceed in the same way as before, once we check that $\dfrac{e^{-\lambda n} P_n^V 1}{u_{n,\lambda}}$ is uniformly bounded above. But

$$u_{n,\lambda}(x) \geq \int_{n-1}^n e^{-\lambda t} P_t^V 1 dt \geq e^{-\lambda n} \int_{n-1}^n e^{-\lambda(t-n)} P_t^V 1 dt$$

and for $n - 1 \leq t \leq n$:

$$[P_n^V 1] = [P_t^V (P_{n-t}^V 1)] \leq e^{\|V\|} [P_t^V 1] \quad .$$

Hence

$$\begin{aligned} u_{n,\lambda}(x) &\geq e^{-\lambda n} e^{-\|V\|} [P_n^V 1](x) \int_{n-1}^n e^{-\lambda(t-n)} dt \\ &\geq e^{-(\|V\|+|\lambda|)} e^{-\lambda n} [P_n^V 1](x) \quad . \end{aligned}$$

□

To complete this section, we see what can be said in case $(\underline{C})$ when

$\{P_t : t > 0\}$ is symmetric with respect to some measure m . To be precise, we will be working under the hypothesis (S.C.) given next.

(S.C.) Let $\{P_x : x \in E\}$ and $\{P_t : t > 0\}$ be as in (C) . Let m be a non-negative Borel measure on E such that $m(K) < \infty$ for all $K \subset\subset E$ and m is concentrated on a σ-compact subset of E . Assume that for all $\phi, \psi \in L^1(m) \cap C_b(E)$ and $t > 0$:

$$\int \phi P_t \psi dm = \int \psi P_t \phi dm \tag{7.28}$$

(7.29) Lemma: Given $V \in C_b(E)$, define $\{P_t^V : t > 0\}$ by (5.14) . Then $\|P_t^V \phi\|_{L^2(m)} \leq \exp(t\|V\|_{C_b(E)})\|\phi\|_{L^2(m)}$, $t > 0$, for all $\phi \in L^1(m) \cap C_b(E)$. In particular, for each $t > 0$, P_t^V admits a unique continuous extension $\overline{P}_t^V$ as an operator from $L^2(m)$ into $L^2(m)$. Moreover, $\overline{P}_t^V$ is self-adjoint for each $t > 0$ and $\{\overline{P}_t^V : t > 0\}$ is a strongly continuous semi-group on $L^2(m)$. Finally, if $\overline{A}_V$ denotes the generator of $\{\overline{P}_t^V : t > 0\}$, then $\mathrm{Dom}(\overline{A}_V) = \mathrm{Dom}(\overline{A}_0)$ and $\overline{A}_V \phi = \overline{A}_0 \phi + V\phi$ for all $\phi \in \mathrm{Dom}(\overline{A}_0)$.

Proof: Note that for all $K \subset\subset E$ and $\phi \in L^1(m) \cap C_b(E)$:

$$\int_K (P_t\phi)^2 dm \leq \int_K P_t\phi^2 dm = \int \phi^2 P_t \chi_K \, dm$$

$$\leq \int \phi^2 dm \ .$$

Hence $\|P_t\phi\|_{L^2(m)} \leq \|\phi\|_{L^2(m)}$, $\phi \in L^1(m) \cap C_b(E)$. Using (5.14) , it is now easy to see that $\|P_t^V\phi\|_{L^2(m)} \leq \exp(t\|V\|_{C_b(E)})\|\phi\|_{L^2(m)}$, $\phi \in L^1(m) \cap C_b(E)$.

Since $L^1(m) \cap C_b(E)$ is dense in $L^2(m)$, we have now established the existence and uniqueness of $\bar{P}_t^V$. Clearly, as an operator on $L^2(m)$, $\bar{P}_t^V$ has norm bounded by $\exp(t\|V\|_{C_b(E)})$. Moreover, it is obvious that $\{\bar{P}_t^V : t > 0\}$ is a semigroup which is weakly continuous on $L^2(m)$. Hence, it is strongly continuous.

To prove that $\bar{P}_t^V$ is self-adjoint, first observe that $\bar{P}_t^0$ is obviously self-adjoint. Next, from (5.14) , it is easily seen that

$$\bar{P}_t^V = \bar{P}_t^0 + \int_0^t \bar{P}_{t-s}^V (V\bar{P}_s^0)ds \quad , \quad t > 0 \quad , \tag{7.30}$$

and therefore that

$$(\bar{P}_t^V)^* = \bar{P}_t^0 + \int_0^t \bar{P}_s^0 (V(\bar{P}_{t-s}^V)^*)ds \quad , \quad t > 0 \quad ,$$

where $(\bar{P}_t^V)^*$ denotes the adjoint of $\bar{P}_t^V$. Note that there is at most one measurable family $\{\bar{Q}_t : t > 0\}$ of operators on $L^2(m)$ satisfying $\|\bar{Q}_t\|_{Hom(L^2(m);L^2(m))} \leq e^{\lambda t}$, $t \geq 0$, for some $\lambda \in [0,\infty)$ and

$$\bar{Q}_t = \bar{P}_t^0 + \int_0^t \bar{P}_s^0 (V\bar{Q}_{t-s})ds \quad , \quad t > 0 \quad . \tag{7.31}$$

We will know that $(\bar{P}_t^V)^* = \bar{P}_t^V$ once we show that $\{\bar{P}_t^V : t > 0\}$ satisfies (7.31) . But to prove this, it is enough to check that

$$P_t^V = P_t + \int_0^t P_s (VP_{t-s}^V)ds \quad , \quad t > 0 \quad . \tag{7.32}$$

But, by (5.15) :

$$P_t^V \phi(x) = E^{P_x}[\exp(\int_0^t V(X(u))du)\phi(X(t))]$$

$$= P_t\phi(x) + E^{P_x}[(\exp(\int_0^t V(X(u))du)-1)\phi(X(t))]$$

$$= P_t\phi(x) + E^{P_x}[\int_0^t V(X(s))\exp(\int_s^t V(X(u))du)\phi(X(t))ds]$$

$$= P_t\phi(x) + \int_0^t E^{P_x}[V(X(s))E^{P_{X(s)}}[\exp(\int_0^{t-s} V(X(u))du)\phi(X(t-s))]]ds$$

$$= P_t\phi(x) + \int_0^t [P_s(VP^V_{t-s}\phi)](x)ds \quad .$$

Hence $\{P^V_t : t > 0\}$ does satisfy (7.32), and so $\overline{P}^V_t$ is self-adjoint.

Finally, from (7.30), note that if $\overline{R}^V_\lambda$ denotes the resolvent of $\overline{A}_V$ for $\lambda > \|V\|_{C_b(E)}$, then $\overline{R}^V_\lambda = \overline{R}^0_\lambda + \overline{R}^V_\lambda \circ (V\overline{R}^0_\lambda)$. Hence $\text{Range}(\overline{R}^0_\lambda) \subseteq \text{Range}(\overline{R}^V_\lambda)$. On the other hand, starting from (7.32), we see that $\overline{R}^V_\lambda = \overline{R}^0_\lambda + \overline{R}^0_\lambda \circ (V\overline{R}^V_\lambda)$; and so we conclude that $\text{Range}(\overline{R}^V_\lambda) = \text{Range}(\overline{R}^0_\lambda)$ for $\lambda > \|V\|_{C_b(E)}$. Since $\text{Dom}(\overline{A}_V) = \text{Range}(\overline{R}^V_\lambda)$ for all $\lambda > \|V\|_{C_b(E)}$, this proves that $\text{Dom}(\overline{A}_V) = \text{Dom}(\overline{A}_0)$. Further, directly from (7.30), it is clear that if $\phi \in \text{Dom}(\overline{A}_0)$, then $\overline{A}_V\phi = \lim_{t\downarrow 0}(\overline{P}^V_t\phi-\phi)/t = \overline{A}_0\phi + V\phi$. □

In the future, we will usually use the notation $\overline{P}_t$ and $\overline{A}$ in place of $\overline{P}^0_t$ and $\overline{A}_0$, respectively.

For $V \in C_b(E)$, define

(7.33) $$\lambda_\sigma(V) = 1/2 \lim_{t\to\infty} \frac{1}{t} \log \|\overline{P}_t^V\|_{\mathrm{Hom}(L^2(m);L^2(m))} \quad .$$

Using $\{E_\sigma^V : \sigma \in R^1\}$ to denote the spectral resolution of $-\overline{A}_V$, we see from

$$\overline{P}_t^V = \int_{R^1} e^{-\sigma t} dE_\sigma^V$$

Noting that $E_\sigma^V = 0$ for $\sigma < -\lambda_\sigma(V)$, we see that

(7.34) $$\overline{P}_t^V = \int_{[-\lambda_\sigma(V),\infty)} e^{-\sigma t} dE_\sigma^V$$

and

(7.35) $$\overline{A}_V = -\int_{[-\lambda_\sigma(V),\infty)} \sigma dE_\sigma^V \quad .$$

In particular,

$$\lambda_\sigma(V) = \sup\{(\phi,\overline{A}_V\phi)_{L^2(m)} : \phi \in D(\overline{A}_V) \text{ and } \|\phi\|_{L^2(m)} = 1\} \quad ,$$

and so

$$\lambda_\sigma(V) = \sup\{\int V(y)\phi(y)^2 m(dy) + (\phi,\overline{A}\ \phi)_{L^2(m)} \ ; \ \phi \in D(\overline{A}) \text{ and } \|\phi\|_{L^2(m)} = 1\} \quad .$$

Next, define the <u>Dirichet form ε</u> by:

(7.36) $$\varepsilon(\phi,\phi) = \int_0^\infty \sigma d(E_\sigma\phi,\phi) \quad , \quad \phi \in L^2(m) \quad .$$

Then, $\varepsilon(\phi,\phi) = -(\phi,\overline{A}\phi)_{L^2(m)}$ for $\phi \in D(\overline{A})$. Moreover, if $\phi \in L^2(m)$, then $\varepsilon(\overline{P}_t\phi,\overline{P}_t\phi) = \int_0^\infty \sigma e^{-2\sigma t} d(E_\sigma\phi,\phi) \uparrow \varepsilon(\phi,\phi)$. Thus

(7.37) $$\lambda_\sigma(V) = \sup\{\int V(y)\phi(y)^2 m(dy) - \varepsilon(\phi,\phi)_{L^2(m)} : \|\phi\|_{L^2(m)} = 1\} \quad .$$

(7.38) Lemma: For each $t > 0$, define

$$\varepsilon_t(\phi,\phi) = \frac{1}{t}(\phi - P_t\phi,\phi)_{L^2(m)} .$$

Then $\varepsilon_t(\phi,\phi) \leq \varepsilon(\phi,\phi)$ and $\varepsilon_t(\phi,\phi) \to \varepsilon(\phi,\phi)$ as $t \downarrow 0$. Moreover, if $m_t(dx \times dy) = P(t,x,dy)m(dx)$ on $E \times E$, then m_t is symmetric (i.e. $m_t(\Gamma_1 \times \Gamma_2) = m_t(\Gamma_2 \times \Gamma_1)$), $m_t(E \times \Gamma) \leq m(\Gamma)$ for all $\Gamma \in \mathcal{B}_E$, and

$$(7.39) \qquad t\varepsilon_t(\phi,\phi) = 1/2\int(\phi(y)-\phi(x))^2 m_t(dx \times dy) + \int(1-\sigma_t(y))\phi(y)^2 m(dy) ,$$

where $\sigma_t(y) = \dfrac{m_t(E\times dy)}{m(dy)}$ can be chosen so that $0 \leq \sigma_t \leq 1$.

Proof: To prove that $\varepsilon(\phi,\phi) = \lim\limits_{t\downarrow 0} \varepsilon_t(\phi,\phi)$, first suppose that $\varepsilon(\phi,\phi) < \infty$. Then, since $(1-e^{-\sigma t})/t \leq \sigma$, $\sigma \geq 0$ and $t > 0$, $\varepsilon(\phi,\phi) = \lim\limits_{t\downarrow 0} \varepsilon_t(\phi,\phi)$ by Lebesgue's dominated convergence theorem. On the other hand, by Fatou's lemma, $\varepsilon(\phi,\phi) \leq \varliminf\limits_{t\downarrow 0} \varepsilon_t(\phi,\phi)$. Thus $\varepsilon(\phi,\phi) = \infty$ implies $\lim\limits_{t\downarrow 0} \varepsilon_t(\phi,\phi) = \infty$.

Next note that for $\phi,\psi \in L^1(m) \cap C_b(E)$:

$$\int\phi(x)\psi(y)m_t(dx \times dy) = \int\phi P_t\psi dm = \int\psi P_t\phi dm = \int\psi(y)\phi(x)m_t(dx \times dy) .$$

Thus m_t is symmetric. Also, if $\phi \in L^1(m) \cap C_b(E)^+$, then

$\int\phi(y)m_t(E \times dy) = \int P_t\phi(x)m(dx) \leq \int\phi(y)m(dy)$. Hence, $m_t(E \times dy) \leq m(dy)$.

Finally, to prove (7.39), first note that both sides are continuous with respect to $L^2(m)$-convergence. Thus, we need only prove (7.39) for $\phi \in L^1(m) \cap C_b(E)$. But then:

$$t\varepsilon_t(\phi,\phi) = \int(\phi-P_t\phi)\phi dm$$

$$= \int\phi^2(y)m(dy) - \int\phi(x)\phi(y)m_t(dx \times dy)$$

$$= \int\phi^2(y)m(dy) + 1/2\int (\phi(x)-\phi(y))^2 m_t(dx \times dy)$$

$$- \int\sigma_t(y)\phi(y)^2 m(dy) \quad .$$

□

(7.40) Lemma: $\phi \to \varepsilon(\phi^{1/2},\phi^{1/2})$ is a l.s.c. convex function on $L^1(m)^+$, and

(7.41) $\lambda_\sigma(V) = \sup\{\int V(y)\phi(y)m(dy) - \varepsilon(\phi^{1/2},\phi^{1/2}) : \phi \in L^1(m)^+ \text{ and } \|\phi\|_{L^1(m)} = 1\}$

Proof: Since $\psi \to \varepsilon(\psi,\psi)$ is l.s.c. on $L^2(m)$, it is clear that $\phi \to \varepsilon(\phi^{1/2},\phi^{1/2})$ is l.s.c. on $L^1(m)$. Moreover, from (7.39), it is clear that $\varepsilon_t(|\psi|,|\psi|) \le \varepsilon_t(\psi,\psi)$ for all $t > 0$ and $\psi \in L^2(m)$. Thus, from (7.37),

$$\lambda_\sigma(V) = \sup\{\int V(y)\psi(y)^2 m(dy) - \varepsilon(\psi,\psi) : \psi \in L^2(m)^+ \text{ and } \|\psi\|_{L^2(m)} = 1\} \quad .$$

Clearly this is equivalent to (7.41). Finally, to prove that $\phi \to \varepsilon(\phi^{1/2},\phi^{1/2})$ is convex on $L^1(m)^+$, it suffices to check that $\phi \to \varepsilon_t(\phi^{1/2},\phi^{1/2})$ is convex on $L^1(m)^+$ for all $t > 0$; and, since $\phi \to \varepsilon_t(f^{1/2},f^{1/2})$ is l.s.c. on $L^1(m)^+$, all we need to do is check that

$$\varepsilon_t\left(\left(\frac{\phi_1+\phi_2}{2}\right)^{1/2},\left(\frac{\phi_1+\phi_2}{2}\right)^{1/2}\right) \le 1/2\left(\varepsilon_t(\phi_1^{1/2},\phi_1^{1/2}) + \varepsilon_t(\phi_2^{1/2},\phi_2^{1/2})\right) \quad .$$ But

$$\varepsilon_t\left(\left(\frac{\phi_1+\phi_2}{2}\right)^{1/2},\left(\frac{\phi_1+\phi_2}{2}\right)^{1/2}\right) = 1/2\varepsilon_t((\phi_1+\phi_2)^{1/2},(\phi_1+\phi_2)^{1/2}) \quad ;$$

and, by triangle inequality for R^2 :

$$t\varepsilon_t((\phi_1 + \phi_2)^{1/2},(\phi_1 + \phi_2)^{1/2})$$

$$= 1/2\int[(\phi_1(y)+\phi_2(y))^{1/2} - (\phi_1(x)+\phi_2(x))^{1/2}]^2 m_t(dx \times dy)$$

$$+ \int(1-\sigma_t(y))(\phi_1(y)+\phi_2(y))m(dy)$$

$$\leq 1/2\int[((\phi_1(x)^{1/2}-\phi_1(y)^{1/2})^2+(\phi_2(x)^{1/2}-\phi_2(y)^{1/2})^2)^{1/2}]^2 m_t(dx \times dy)$$

$$+ \int(1-\sigma_t(y))\phi_1(y)m(dy) + \int(1-\sigma_t(y))\phi_2(y)m(dy)$$

$$= t\varepsilon_t(\phi_1^{1/2},\phi_1^{1/2}) + t\varepsilon_t(\phi_2^{1/2},\phi_2^{1/2}) \quad . \qquad \square$$

Now define $J_\sigma : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ by

(7.42) $$J_\sigma(\nu) = \begin{cases} \infty & \text{if } \nu \notin \mathcal{M}_1(E) \text{ or } \nu \in \mathcal{M}_1(E) \text{ but } \nu \not\ll m \\ \varepsilon\left(\left(\frac{d\nu}{dm}\right)^{1/2}, \left(\frac{d\nu}{dm}\right)^{1/2}\right) & \text{if } \nu \in \mathcal{M}_1(E) \text{ and } \nu \ll m \end{cases}$$

Clearly, by Lemma (7.40), $\nu \to J_\sigma(\nu)$ is convex and

(7.43) $$\lambda_\sigma = J_\sigma^* \quad .$$

If we knew that J_σ were l.s.c. on $\mathcal{M}_1(E)$, then we could say that $J_\sigma = \lambda_\sigma^*$. What we are going to show is that, under an additional hypothesis, not only is J_σ l.s.c. but, in fact, $J_\sigma = J$. We will then have that $\lambda_\sigma^* = J_\sigma = J = \lambda^*$.

(7.44) Theorem: For all $\nu \in \mathcal{M}_1(E)$, $J(\nu) \leq J_\sigma(\nu)$. If $\nu \in \mathcal{M}_1(E)$ and $\nu P_h \ll m$ for all $h > 0$, then $J_\sigma(\nu) = J(\nu)$. Thus, if $\nu P_h \ll m$ for all $\nu \in \mathcal{M}_1(E)$ and $h > 0$, then $J_\sigma = J$ and so $\lambda_\sigma^* = J_\sigma = J = \lambda^*$.

Proof: To see that $J(\nu) \leq J_\sigma(\nu)$ for all ν, we will assume that $\nu \in \mathcal{M}_1(E)$, $\nu \ll m$, and $\varepsilon(f^{1/2},f^{1/2}) < \infty$ where $f = \frac{d\nu}{dm}$. Given $u \in \mathcal{U} \cap D(A)$

define $V_u = -\frac{Au}{u}$. Then $\overline{P}_t^{V_u} = \int_{-\lambda_u}^{\infty} e^{-\sigma t} dE_\sigma^u$, where $\lambda_u = \lambda_\sigma(V_u)$ and $\{E_\sigma^u : \sigma \geq -\lambda_u\}$ is the spectral resolution of $-\overline{A}_{V_u}$. If $\lambda < \lambda_u$, we know that $E_{-\lambda}^u - E_{-\lambda_u}^u \neq 0$. Thus there is a $\phi \in L^1(m) \cap C_b(E)$ such that $(E_{-\lambda}^u - E_{-\lambda_u}^u)\phi \neq 0$. In particular, $\underline{\lim}_{t\to\infty} \frac{1}{t} \log(\phi, \overline{P}_t^{V_u}\phi)_{L^2(m)} \geq \lambda$. At the same time, we have already seen that $\lambda(V_u) = 0$ (cf. the proof of (7.12)); and clearly

$$(\phi, \overline{P}_t^{V_u}\phi)_{L^2(m)} \leq \|\phi\|_{L^1(m)} \|\phi\|_{C_b(E)} \|P_t^{V_u}\|_{op} .$$

Thus, $\lambda \leq 0$. We now know that $\lambda_u \leq 0$. But this means that $\overline{A}_{V_u} \leq 0$; and so:

$$(\psi, \overline{A}_{V_u} \psi)_{L^2(m)} \leq 0$$

for all $\psi \in D(\overline{A})$. In other words, we now see that:

$$\varepsilon(\psi,\psi) \geq -\int \frac{Au}{u} \psi^2 dm$$

for all $\psi \in D(\overline{A})$ and $u \in \mathcal{U}$. Choosing $\{\psi_n\}_1^\infty \subseteq D(\overline{A})$ so that $\psi_n \to f^{1/2}$ in $L^2(m)$ and $\varepsilon(\psi_n,\psi_n) \to (f^{1/2}, f^{1/2})$ (e.g.. $\psi_n = \overline{P}_{1/n} f^{1/2}$) , we conclude that $\varepsilon(f^{1/2}, f^{1/2}) \geq \int -\frac{Au}{u} d\nu$ for all $u \in \mathcal{U} \cap D(A)$.

Next, assume that $\nu \in \mathcal{M}_1(E)$, $\ell \equiv J(\nu) < \infty$, and that $\nu P_h \ll m$ for all $h > 0$. Since, by Theorem (7.25) , $\|\nu P_h - \nu\|_{var} \to 0$ as $h \downarrow 0$, we see that $\nu \ll m$. Set $f = \frac{d\nu}{dm}$. By Corollary (7.24) :

$$\int \log \frac{P_h u}{u}\, d\nu \geq -h\ell$$

for all $u \in \mathcal{U}$. Since $\log(1-x) \leq -x$ for $x \in (-\infty,1]$ and because $\frac{u-P_h u}{u} \leq 1$, we have that

$$-\int \frac{u-P_h u}{u}\, d\nu \geq \int \log \frac{P_h u}{u}\, d\nu \geq -h\ell$$

for all $u \in \mathcal{U}$. In other words,

$$\sup_{u\in\mathcal{U}} \int \frac{u-P_h u}{u}\, fdm \leq h\ell \quad .$$

Given $n \geq 1$ and $\varepsilon > 0$, choose $\{u_k\} \subseteq C_b(E)$ so that $\varepsilon \leq u_k \leq n + \varepsilon$ and $u_k \to f^{1/2} \wedge n + \varepsilon$ in m-measure. Then we get, from the above, that

$$\int \frac{f^{1/2} \wedge n - \overline{P}_h(f^{1/2} \wedge n)}{f^{1/2} \wedge n + \varepsilon}\, fdm \leq h\ell$$

for all $n \geq 1$ and $\varepsilon > 0$. Note that

$$0 \leq \frac{f^{1/2} \wedge n}{f^{1/2} \wedge n + \varepsilon}\, f \leq f \in L^1(m)$$

and

$$0 \leq \frac{\overline{P}_h(f^{1/2} \wedge n)}{f^{1/2} \wedge n + \varepsilon}\, f \leq ((\overline{P}_h f^{1/2}) \wedge n)(f^{1/2} \vee \frac{f}{n})$$

$$\leq (f^{1/2} \overline{P}_h f^{1/2}) \vee f \in L^1(m)$$

for all $n \geq 1$ and $\varepsilon > 0$. Thus, by Lebesgue's dominated convergence theorem, we conclude that:

$$\int (f^{1/2} - \overline{P}_h f^{1/2}) f^{1/2} dm \leq h\ell \quad .$$

Hence $\varepsilon(f^{1/2},f^{1/2}) = \lim_{h\downarrow 0} \varepsilon_h(f^{1/2},f^{1/2}) \leq \ell$.

8. Some Non-Uniform Large Deviation Results:

Thus far we have a quite satisfactory theory of large deviations for Markov processes under the assumption (6.1) , in case $(\underline{D})$, or (6.8) , in case $(\underline{C})$. Unfortunately, when E is not compact, such an assumption will not be satisfied, except in very special situations (e.g. the Sanov theorem). For example, let $E = R^1$ and $P(t,x,dy) = g(1-e^{-2t},y-e^{-t}x)dy$, where $g(s,\xi) = (2\pi s)^{-1/2}\exp(-\xi^2/2s)$. Clearly, the associated process (i.e. the Ornstein-Uhlenbeck process) has strong ergodic properties and one ought to be able to study the large deviation theory. At the same time, it is equally clear that $P(t,x,\cdot)$ fails to satisfy (6.8) , and therefore the theory developed in section 6) is not applicable. Of course, the reason why section 6) cannot handle this process is clear; namely: when section 6) applies, the resulting large deviation principle is uniform, whereas one should not expect a uniform principle to hold for the Ornstein-Uhlenbeck process. Indeed, one can hope for uniform large deviation principles only in the presence of uniform ergodicity.

There are several ways in which one might try to handle situations in which (6.1) or (6.8) fail to hold. One approach would be to attempt a variation on the theme of section 6) and seek a convex rate function $I : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ such that, in case $(\underline{D})$:

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log Q_{n,x}(F) \leq -\inf_{\nu\in F} I(\nu) , \quad F \text{ closed,}$$

$$\underline{\lim_{n\to\infty}} \frac{1}{n} \log Q_{n,x}(G) \geq -\inf_{\nu\in G} I(\nu) , \quad G \text{ open,}$$

for all $x \in E$; and the analogous result in case $(\underline{C})$. As a consequence of Theorem (2.6) , one would then have that

$$\tilde{\lambda}(V) = I^{*}(V) \ , \quad V \in C_b(E) \quad ;$$

and therefore, by theorem (7.27) , one would know that

$$I = J \quad .$$

Unfortunately, at the present time, the details of such an approach have not been worked through. Thus, we will adopt a different tack, more along the lines of the one used by Donsker and Varadhan in their original paper. This approach will yield quite satisfactory results for the upper bound (i.e. for closed sets), but will give us a lower bound only for symmetric processes.

As we have just pointed out, Theorem (7.27) tells us that if there is a convex rate function I , it pretty much has to be J . Thus, it is reasonable to see how well one can do by just guessing that J is the right rate function and seeing if one can make it work. The procedure which we are about to use is essentially the same as the one with which we proved the original Cramér theorem (cf. Theorem (3.8)). We begin with a very general result.

(8.1) <u>Theorem</u>: Let C be a compact subset of $\mathcal{M}_1(E)$. Then, in case $(\underline{D})$:

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(C)) \leq -\inf_{\nu\in C} J(\nu) \quad ;$$

and, in case $(\underline{C})$:

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(C)) \leq -\inf_{\nu\in C} J(\nu) \quad .$$

Proof: Set $\ell = \inf_{\nu \in C} J(\nu)$. Given $\varepsilon > 0$, choose, for each $\nu \in C$, a $V_\nu \in C_b(E)$ so that

$$\int V_\nu d\nu - \lambda(V_\nu) \geq \ell - \varepsilon \quad .$$

(This can be done by Theorem (7.18) .) Next, for each $\nu \in C$, choose an open neighborhood B_ν of ν so that

$$\sup_{\mu \in B_\nu} \left| \int V_\nu d\mu - \int V_\nu d\nu \right| < \varepsilon \quad .$$

Select a finite number N of $\nu_1, \ldots, \nu_N \in C$ so that $C \subseteq \bigcup_1^N B_{\nu_i}$. Clearly in case (D) :

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(C)) \leq \max_{1\leq i\leq N} \overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in E} Q_{n,x}(B_{\nu_i})) \quad ;$$

and, in case (C) :

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(C)) \leq \max_{1\leq i\leq N} \overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in E} Q_{t,x}(B_{\nu_i})) \quad .$$

But for $\nu \in C$:

$$\begin{aligned}
\sup_{x\in E} Q_{n,x}(B_\nu) &\leq \sup_{x\in E} E^{P_x}[\exp(\sum_0^{n-1} V_\nu(X(m))) \quad , \quad L_n \in B_\nu] \\
&\quad \times \sup_{\mu\in B_\nu} \exp(-n\int V_\nu(y)\mu(dy)) \\
&\leq \sup_{x\in E} E^{P_x}[\exp(\sum_0^{n-1} V_\nu(X(m)))]\exp(-n\int V_\nu(y)\nu(dy) + n\varepsilon) \\
&= \|\pi^n_{V_\nu}\|_{op}\exp(-n\int V_\nu(y)\nu(dy) + n\varepsilon) \ .
\end{aligned}$$

Thus

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log \sup_{x\in E} Q_{n,x}(B_\nu) \leq \lambda(V_\nu) - \int V_\nu(y)\nu(dy) + \varepsilon$$

$$\leq -\ell + 2\varepsilon \quad .$$

A similar argument applies in case (C) .

□

In order to remove the restriction to compact subsets of $\mathcal{M}_1(E)$, we need to impose some more conditions.

Given a map $\Lambda : C_b(E) \to R^1$, we will say that Λ is tight if for each $M < \infty$ there is a $K(M) \subset\subset E$ such that $|\Lambda(V)| \leq 1$ whenever $V \in C_b(E)$ vanishes on $K(M)$ and satisfies $\|V\|_{C_b(E)} \leq M$.

In the following results, we will be using some new notation. Namely, ρ denotes a topologically equivalent metric on E with the property that (E,ρ) is totally bounded. $\hat{E}$ denotes the completion of E with respect to ρ . Clearly $\hat{E}$ is compact. Moreover, because E is Polish, E is a dense G_δ subset of $\hat{E}$. Denote by $\hat{C}(E)$ the space of ρ-uniformly continuous functions of E into R^1 . Then $\hat{V} \in C(\hat{E}) \to \hat{V}|_E$ is an isomorphism from $C(\hat{E})$ onto $\hat{C}(E)$. Finally, if $\hat{\nu} \in \mathcal{M}_1(\hat{E})$, then, because $E \in \mathcal{B}_{\hat{E}}$, $\hat{\nu} = \theta\nu_1 + (1-\theta)\nu_2$ where $\theta \in [0,1]$, $\nu_1 \in \mathcal{M}_1(E)$, and $\nu_2 \in \mathcal{M}_1(\hat{E})$ with $\nu_2(E) = 0$.

(8.2) Lemma: Let $\Lambda : C_b(E) \to R^1$ be a tight convex function satisfying $\Lambda(V_1) \leq \Lambda(V_2)$ whenever $V_1, V_2 \in C_b(E)$ satisfy $V_1 \leq V_2$. Further, assume that $\Lambda(c1) = c$ for all $c \in R^1$. Then Λ is continuous. Moreover, for each $\varepsilon > 0$ and $0 < M < \infty$, there is a $K(\varepsilon,M) \subset\subset E$ such that $|\Lambda(V_1)-\Lambda(V_2)| < \varepsilon$ whenever $V_1,V_2 \in C_b(E)$ satisfy $V_1 = V_2$ on $K(\varepsilon,M)$ and

$\|V_1\|_{C_b(E)} \vee \|V_2\|_{C_b(E)} \le M$. Next, define

$$I(\nu) = \Lambda^*(\nu) \equiv \sup\{\int V d\nu - \Lambda(V) : V \in C_b(E)\}$$

for $\nu \in \mathcal{M}(E)$ and

$$\hat{I}(\hat{\nu}) = \sup\{\int \hat{V} d\hat{\nu} - \Lambda(\hat{V}|_E) : \hat{V} \in C(\hat{E})\}$$

for $\hat{\nu} \in \mathcal{M}(\hat{E})$. Then, $\{\hat{\nu} \in \mathcal{M}(\hat{E}) : \hat{I}(\hat{\nu}) \le L\} \subset\subset \mathcal{M}_1(E)$ for all $L > 0$ and $\hat{I}(\nu) = I(\nu)$ for $\nu \in \mathcal{M}(E)$. Finally, $\Lambda(V) = I^*(V)$ for $V \in \hat{C}(E)$ and there is a $\nu_0 \in \mathcal{M}_1(E)$ such that $I(\nu_0) = 0$.

<u>Proof</u>: Let $V_1, V_2 \in C_b(E)$ be given. For $\theta \in (0,1)$, $V_1 = (1-\theta)V_2 + \theta(\frac{V_1-(1-\theta)V_2}{\theta})$. Thus,

$$\Lambda(V_1) \le (1-\theta)\Lambda(V_2) + \theta\Lambda(\frac{V_1-(1-\theta)V_2}{\theta})$$

$$\le (1-\theta)\Lambda(V_2) + \frac{\theta}{2}(\Lambda(\frac{2(V_1-V_2)}{\theta}) + \Lambda(2V_2)) \quad .$$

Thus, for all $\theta \in (0,1)$ and $V_1, V_2 \in C_b(E)$:

$$(8.3) \qquad \Lambda(V_1) - \Lambda(V_2) \le \frac{\theta}{2}\,\Lambda(\frac{2(V_1-V_2)}{\theta}) + \theta(1/2\Lambda(2V_2) - \Lambda(V_2)) \quad .$$

Note that for all $V \in C_b(E)$, $\Lambda(V) \le \Lambda(\|V\|_{C_b(E)}1) = \|V\|_{C_b(E)}$. Also, for $V \in C_b(E)$, $0 = \Lambda(01) = \Lambda(\frac{V-V}{2}) \le (\Lambda(V)+\Lambda(-V))/2$; and so, $-\Lambda(V) \le \|V\|_{C_b(E)}$. Thus, $|\Lambda(V)| \le \|V\|_{C_b(E)}$. Using this in (8.3) , we

$$\Lambda(V_1) - \Lambda(V_2) \le \|V_1-V_2\|_{C_b(E)} + 2\theta\|V_2\|_{C_b(E)}$$

for all $\theta \in (0,1)$. Letting $\theta\downarrow 0$ and exchanging V_1 with V_2 , we

conclude that

$$(8.4) \qquad \left|\Lambda(V_1) - \Lambda(V_2)\right| \le \|V_1-V_2\|_{C_b(E)} \quad .$$

This certainly proves that Λ is continuous.

Returning to (8.3) and using the preceding, we see that

$$\Lambda(V_1) - \Lambda(V_2) \le \frac{\theta}{2}(\Lambda(\frac{2(V_1-V_2)}{\theta}) + 4\|V_2\|_{C_b(E)})$$

for any $\theta \in (0,1)$ and $V_1,V_2 \in C_b(E)$. Hence

$$\left|\Lambda(V_1)-\Lambda(V_2)\right| \le \frac{\theta}{2}(\Lambda(\frac{2\left|V_1-V_2\right|}{\theta}) + 4(\|V_1\|_{C_b(E)} \vee \|V_2\|_{C_b(E)})) \; .$$

Given $\varepsilon > 0$ and $0 < M < \infty$, take $\theta = \frac{\varepsilon}{4M+1}$ and $K(\varepsilon,M) = K(4M/\theta)$. If $V_1 = V_2$ on $K(\varepsilon,M)$ and $\|V_1\|_{C_b(E)} \vee \|V_2\|_{C_b(E)} \le M$, then we have:

$$\left|\Lambda(V_1)-\Lambda(V_2)\right| \le \frac{\varepsilon}{8M+2}(1+4M) = \frac{\varepsilon}{2} < \varepsilon \quad .$$

Since $\Lambda(c1) = c$ for $c \in R^1$, it is clear that $I(\nu) = \infty$ and $\hat{I}(\hat{\nu}) = \infty$ unless $\nu \in \mathcal{M}_1(E)$ and $\hat{\nu} \in \mathcal{M}_1(\hat{E})$, respectively. Next, suppose that $\hat{\nu} \in \mathcal{M}_1(\hat{E})\setminus \mathcal{M}_1(E)$. Then $\hat{\nu} = (1-\theta)\nu_1 + \theta\nu_2$ for some $\theta \in (0,1]$, $\nu_1 \in \mathcal{M}_1(E)$ and $\nu_2 \in \mathcal{M}_1(\hat{E})$ with $\nu_2(E) = 0$. Since E is a G_δ in $\hat{E}$, $\hat{E}\setminus E$ is the countable union of compact sets. Thus, we can choose $\hat{K} \subset\subset \hat{E}\setminus E$ so that $\nu_2(\hat{K}) > 0$. Given $0 < M < \infty$, we use the Tietze extension theorem to construct a $\hat{V}_M \in C(\hat{E})$ and that $\hat{V}_M = M$ on $\hat{K}$, $\hat{V}_M = 0$ on $K(M)$ and $0 \le \hat{V}_M \le M$ on $\hat{E}$. Then:

$$\begin{aligned}\hat{I}(\hat{\nu}) &\ge \int \hat{V}_M d\hat{\nu} - \Lambda(\hat{V}_M|_E) \\ &\ge \theta M \nu_2(\hat{K}) - 1 \quad .\end{aligned}$$

Letting $M\uparrow\infty$, we conclude that $\hat{I}(\hat{\nu}) = \infty$.

We now know that $\hat{I}(\hat{\nu}) = \infty$ unless $\hat{\nu} \in \mathcal{M}_1(E)$ and that $I(\nu) = \infty$ unless $\nu \in \mathcal{M}_1(E)$. For $\nu \in \mathcal{M}_1(E)$, it is clear that $\hat{I}(\nu) \leq I(\nu)$. Next, suppose that $\nu \in \mathcal{M}_1(E)$ and that $I(\nu) < \infty$. Given $\varepsilon > 0$, choose $V \in C_b(E)$ so that

$$\int V d\nu - \Lambda(V) \geq I(\nu) - \varepsilon \quad .$$

Set $M = \|V\|_{C_b(E)}$ and choose $K \subset\subset E$ so that $K(\varepsilon,M) \subseteq K$ and $\nu(K^c) \leq \varepsilon/M$. Use the Tietze extension theorem to construct a $\hat{V} \in \hat{C}(E)$ so that $\hat{V} = V$ on K and $\|\hat{V}\|_{\hat{C}(\hat{E})} \leq M$. Then $\left|\Lambda(\hat{V}|_E) - \Lambda(V)\right| \vee \left|\int V d\nu - \int \hat{V} d\nu\right| \leq \varepsilon$. Hence

$$\hat{I}(\nu) \geq \int \hat{V} d\nu - \Lambda(\hat{V}|_E) \geq I(\nu) - 3\varepsilon \quad .$$

A similar argument applies to the case when $I(\nu) = \infty$. Thus, $I = \hat{I}$ on $\mathcal{M}_1(E)$; and so $I = \hat{I}$ on $\mathcal{M}(E)$.

We next show that $\{\hat{\nu} \in \mathcal{M}_1(\hat{E}) : \hat{I}(\hat{\nu}) \leq L\} \subset\subset \mathcal{M}_1(E)$ for all $L > 0$. In view of the preceding, this reduces to checking that $\{\nu \in \mathcal{M}_1(E) : I(\nu) \leq L\} \subset\subset \mathcal{M}_1(E)$ for all $L > 0$. But if $\nu \in \mathcal{M}_1(E)$ and $I(\nu) \leq L$, then for all $0 < M < \infty$ and $V \in C_b(E)$ satisfying $V = 0$ on $K(M)$ and $\|V\|_{C_b(E)} \leq M$:

$$\int V d\nu - 1 \leq \int V d\nu - \Lambda(V) \leq L \quad ;$$

and so $\nu(K(M)^c) \leq \frac{L+1}{M}$. In other words, $\{\nu \in \mathcal{M}_1(E) : I(\nu) \leq L\}$ is contained in the compact set

$$\{\nu \in \mathcal{M}_1(E) : \nu(K(M)^c) \leq \frac{L+1}{M} \text{ for all } 0 < M < \infty\} \quad .$$

Since I is l.s.c., this proves that $\{\nu \in \mathcal{M}_1(E) : I(\nu) \leq L\}$ is compact.

We finally show that $I^*(V) = \Lambda(V)$, $V \in \hat{C}(E)$. From our knowledge of I

and $\hat{I}$, it is clear that for $V \in \hat{C}(E)$: $I^*(V) = (\hat{I})^*(\hat{V}) \equiv \sup\{\int \hat{V} d\hat{\nu} - \hat{I}(\hat{\nu}) : \hat{\nu} \in \mathcal{M}_1(\hat{E})\}$, where $\hat{V}$ is the unique element of $C(\hat{E})$ satisfying $V = \hat{V}|_E$. At the same time, $\hat{I} = (\hat{\Lambda})^*$, where $\hat{\Lambda}(\hat{V}) = \Lambda(\hat{V}|_E)$ for $\hat{V} \in C(\hat{E})$. Since $\mathcal{M}(\hat{E})$ is the dual of $C_b(\hat{E})$, Theorem (7.15) tells us that $\hat{\Lambda} = (\hat{I})^*$. Combining these, we see that $\Lambda(V) = I^*(V)$ for $V \in \hat{C}(E)$.

The existence of ν_0 such that $I(\nu_0) = 0$ is derived as follows. Since $1 = \Lambda(1) = I^*(1)$, for each $n \geq 1$ we can find a $\nu_n \in \mathcal{M}_1(E)$ such that $1 - I(\nu_n) = \int 1 d\nu_n - I(\nu_n) \geq 1 - 1/n$. Hence $I(\nu_n) \leq 1/n$, $n \geq 1$. In particular, $\{\nu_n\}_1^\infty \subseteq \{I \leq 1\} \subset\subset \mathcal{M}_1(E)$. Let ν_0 be any limit of $\{\nu_n\}_1^\infty$. Since I is l.s.c., $I(\nu_0) = 0$. □

(8.5) <u>Lemma</u>: Let $I : \mathcal{M}(E) \to [0,\infty) \cup \{\infty\}$ be a l.s.c. function satisfying $\{\nu \in \mathcal{M}(E) : I(\nu) \leq L\} \subset\subset \mathcal{M}_1(E)$ for all $L > 0$ and for which there is a $\nu_0 \in \mathcal{M}_1(E)$ such that $I(\nu_0) = 0$. Then $I^* : C_b(E) \to R^1$ satisfies all the hypotheses of Lemma (8.2). In particular, if $\Lambda : C_b(E) \to R^1$ satisfies the hypotheses of Lemma (8.2) and $I = \Lambda^*$, then $\Lambda = I^*$ on the whole of $C_b(E)$.

<u>Proof</u>: First, note that if $V_1 \leq V_2$, then

$$\begin{aligned} I^*(V_1) &= \sup\{\int V_1 d\nu - I(\nu) : \nu \in \mathcal{M}_1(E)\} \\ &\leq \sup\{\int V_2 d\nu - I(\nu) : \nu \in \mathcal{M}_1(E)\} \\ &= I^*(V_2) . \end{aligned}$$

Thus $I^*(V_1) \leq I^*(V_2)$ if $V_1 \leq V_2$. Next,

$$I^*(c1) = \sup\{c - I(\nu) : \nu \in \mathcal{M}_1(E)\} = c \quad ,$$

since $I(\nu_0) = 0$ and $I(\nu) \geq 0$ for all $\nu \in \mathcal{M}_1(E)$.

To prove that I^* is tight, let $0 < M < \infty$ be given and set $C = \{I \leq 2M\}$. Then $C \subset\subset \mathcal{M}_1(E)$ and so we can find $K(M) \subset\subset E$ so that $\nu(K(M))^c < 1/M$ for $\nu \in C$. Given $V \in C_b(E)$ satisfying $V = 0$ on $K(M)$ and $\|V\|_{C_b(E)} \leq M$, we have:

$$\begin{aligned} I^*(V) &= \sup\{\textstyle\int V d\nu - I(\nu) : \nu \in C\} \\ &\qquad \vee \sup\{\textstyle\int V d\nu - I(\nu) : \nu \in \mathcal{M}_1(E)\setminus C\} \\ &\leq \sup\{M\nu((K(M))^c) : \nu \in C\} \vee (-M) \\ &\leq 1 \quad . \end{aligned}$$

At the same time, $-I^*(V) \leq I^*(-V)$ for all $V \in C_b(E)$. Hence $|I^*(V)| \leq 1$ for all $V \in C_b(E)$ satisfying $V = 0$ on $K(M)$ and $\|V\|_{C_b(E)} \leq M$.

Now suppose that Λ is as in Lemma (8.2) and that $I = \Lambda^*$. By Lemma (8.2) , I satisfies the hypotheses of this lemma. Thus, by what we have just seen, Lemma (8.2) applies to I^* as well as Λ . In particular, for each $\varepsilon > 0$ and $0 < M < \infty$, there is a $K(\varepsilon,M) \subset\subset E$ such that $|I^*(V_1) - I^*(V_2)| \vee |\Lambda(V_1) - \Lambda(V_2)| < \varepsilon$ for all $V_1, V_2 \in C_b(E)$ satisfying $V_1 = V_2$ on $K(\varepsilon,M)$ and $\|V_1\|_{C_b(E)} \vee \|V_2\|_{C_b(E)} \leq M$. Now choose any $V \in C_b(E)$. Set $M = \|V\|_{C_b(E)}$ and for each $n \geq 1$ use the Tietze extension theorem to construct a $\hat{V}_n \in C(\hat{E})$ so that $\hat{V}_n = V$ on $K(1/n,M)$ and $\|\hat{V}_n\|_{C(\hat{E})} \leq M$. Set $V_n = \hat{V}_n|_E$. By Lemma (8.2) , $I^*(V_n) = \Lambda(V_n)$, $n \geq 1$. On the other hand, $|I^*(V_n) - I^*(V)| \vee |\Lambda(V_n) - \Lambda(V)| \leq 1/n$. Hence

$I^*(V) = \Lambda(V)$. □

(8.6) Theorem: Assume that $\tilde{\lambda}$ is tight. Then: J is a rate function, there is a $\nu_0 \in \mathcal{M}_1(E)$ for which $J(\nu_0) = 0$, and $\tilde{\lambda} = J^*$. Moreover, in case (D) :

$$\sup_{x\in E} \overline{\lim_{n\to\infty}} \frac{1}{n} \log Q_{n,x}(F) \leq -\inf_{\nu\in F} J(\nu)$$

and in case (C) :

$$\sup_{x\in E} \overline{\lim_{t\to\infty}} \frac{1}{t} \log Q_{t,x}(F) \leq -\inf_{\nu\in F} J(\nu)$$

for all closed $F \subseteq \mathcal{M}_1(E)$.

Proof: Since, if it is tight, $\tilde{\lambda}$ satisfies the conditions of Lemma (8.2) and because, by Theorem (7.27) , $J = (\tilde{\lambda})^*$, the first part of this theorem follows immediately from Lemma (8.2) and the second part of Lemma (8.5) .

To prove the second part, define $\hat{J}(\hat{\nu}) = \sup\{\int \hat{V} d\hat{\nu} - \hat{\lambda}(\hat{V}|_E) : \hat{V} \in C(\hat{E})\}$ for $\hat{\nu} \in \mathcal{M}(\hat{E})$. Given a closed $F \subseteq \mathcal{M}_1(E)$, let $\hat{F}$ be the closure of F in $\mathcal{M}(\hat{E})$. Then $\hat{F} \subset\subset \mathcal{M}_1(\hat{E})$. Repeating the same argument given to prove Theorem (8.1) (only this time, do the covering in $\hat{E}$ and use $\hat{V} \in C(\hat{E})$), we arrive at the desired inequalities with $\inf_{\nu\in F} J(\nu)$ replaced by $\inf_{\hat{\nu}\in\hat{F}} \hat{J}(\hat{\nu})$ on the right hand side. Finally, by Lemma (8.2) , $\inf_{\hat{\nu}\in\hat{F}} \hat{J}(\hat{\nu}) = \inf_{\hat{\nu}\in\hat{F}\cap\mathcal{M}(E)} J(\nu)$; and therefore, since $F = \hat{F} \cap \mathcal{M}_1(E)$, $\inf_{\hat{\nu}\in\hat{F}} \hat{J}(\hat{\nu}) = \inf_{\nu\in F} J(\nu)$. □

In order to develop a criterion which guarantees that $\tilde{\lambda}$ is tight and, at the same time, tells us when we can make the estimates in Theorem (8.6) uniform (at least locally), we need the following simple lemma.

(8.7) Lemma: Let $\Phi : E \to R^1$ be a function which is bounded below and has the property that for each $L > 0$ the set $\{x \in E : \Phi(x) \leq L\}$ is compact. Let $\Gamma \subseteq E$. In case $(\underline{D})$, the condition

$$\sup_{n\geq 1} \sup_{x\in\Gamma} E^{P_x}[\exp(\sum_0^{n-1} \Phi(X(m)))] < \infty \tag{8.8}$$

guarantees that for each $L > 0$ there exists a $C(L) \subset\subset \mathcal{M}_1(E)$ such that:

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x\in\Gamma} Q_{n,x}(C(L)^c)) < -L \quad . \tag{8.9}$$

In case $(\underline{C})$, the condition

$$\sup_{t>0} \sup_{x\in\Gamma} E^{P_x}[\exp(\int_0^t \Phi(X(s))ds)] < \infty \tag{8.10}$$

guarantees that for each $L > 0$ there exists a $C(L) \subset\subset \mathcal{M}_1(E)$ such that:

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x\in\Gamma} Q_{t,x}(C(L)^c)) \leq -L \quad . \tag{8.11}$$

Proof: We work in case $(\underline{D})$, the case $(\underline{C})$ being essentially the same. Choose $L_0 \in (0,\infty)$ so that $\Phi \geq -L_0$ and set $\Psi = \Phi + L_0$. Then

$$E^{P_x}[\exp(\sum_0^{n-1} \Psi(X(m)))] = e^{nL_0} E^{P_x}[\exp(\sum_0^{n-1} \Phi(X(m)))] \quad ,$$

and so there is a $C < \infty$ such that

$$\sup_{x\in\Gamma} E^{P_x}[\exp(\sum_0^{n-1} \Psi(X(m)))] \leq Ce^{nL_0} \quad , \quad n \geq 1 \quad .$$

Given $M \in (0,\infty)$, set $K(M) = \{x \in E : \Psi(x) \leq M^2\}$. Then $K(M) \subset\subset E$ and

$$\sup_{x\in\Gamma} Q_{n,x}(\nu : \nu(K(M)^c) \geq \frac{1}{M})$$

$$\leq \sup_{x \in \Gamma} e^{-nM_E} P_x[\exp(\sum_0^{n-1} \Psi(X(m)))]$$

$$\leq Ce^{-n(M-L_0)} \quad , \quad n \geq 1 \quad .$$

Hence, if $L \in (0,\infty)$ and $C(L) = \bigcap_{\ell=0}^{\infty} \{\nu : \nu(K(\ell+L+L_0)^c) \leq \frac{1}{\ell+L+L_0}\}$, then $C(L) \subset\subset \mathcal{M}_1(E)$ and

$$\sup_{x \in \Gamma} Q_{n,x}(C(L)^c)$$

$$\leq \sum_{\ell=0}^{\infty} \sup_{x \in \Gamma} Q_{n,x}(\{\nu : \nu(K(\ell+L+L_0)^c) \geq \frac{1}{\ell+L+L_0}\})$$

$$\leq Ce^{-nL} \sum_{\ell=0}^{\infty} e^{-n\ell} = \frac{C}{1-e^{-1}} e^{-nL} \quad , \quad n \geq 1 \quad .$$

□

(8.12) <u>Theorem</u>: Let Φ be as in Lemma (8.7) . In case ($\underline{D}$) , condition (8.8) guarantees that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log(\sup_{x \in \Gamma} Q_{n,x}(F)) \leq -\inf_{\nu \in F} J(\nu) \tag{8.13}$$

for all closed F in $\mathcal{M}_1(E)$; and in case ($\underline{C}$) , condition (8.10) guarantees that

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log(\sup_{x \in \Gamma} Q_{t,x}(F)) \leq -\inf_{\nu \in F} J(\nu) \tag{8.14}$$

for all closed F in $\mathcal{M}_1(E)$. Finally, in case ($\underline{D}$) , if, for every $x \in E$, (8.8) holds with $\Gamma = \{x\}$, then $\tilde{\lambda}$ is tight, and, in case ($\underline{C}$) , if, for every $x \in E$, (8.10) holds with $\Gamma = \{x\}$, then $\tilde{\lambda}$ is tight. In particular, in either of these cases, the conclusions of Theorem (8.6) hold.

<u>Proof</u>: Again, we need only work in case ($\underline{D}$) .

Assuming that (8.8) holds, and given a closed F in $\mathcal{M}_1(E)$, for each $L > 0$ choose $C(L) \subset\subset \mathcal{M}_1(E)$ so that (8.9) holds. Then

$$\log(\sup_{x\in\Gamma} Q_{n,x}(F)) \leq \log(2[\sup_{x\in\Gamma} Q_{n,x}(F\cap C(L)) \vee \sup_{x\in\Gamma} Q_{n,x}(C(L)^c)]) ,$$

and so, by (8.8) and Theorem (8.1) :

$$\begin{aligned} &\overline{\lim_{n\to\infty}} \frac{1}{n} \log \sup_{x\in\Gamma} Q_{n,x}(F) \\ &\leq -(\inf_{\nu\in F\cap C(L)} J(\nu) \wedge L) \\ &\leq -(\inf_{\nu\in F} J(\nu) \wedge L) . \end{aligned}$$

Letting $L\uparrow\infty$, we get (8.13) .

Next, suppose that (8.8) holds when $\Gamma = \{x\}$ for every $x \in E$. Given $x\in E$ and $M \in (0,\infty)$, choose $C(L) \subset\subset \mathcal{M}_1(E)$ so that (8.9) holds when $\Gamma = \{x\}$ and $L = M+1$. Choose $K(M) \subset\subset E$ so that $\nu(K(M)^c) \leq 1/M$ for $\nu \in C(L)$. Then:

$$\begin{aligned} &E^{P_x}[\exp(M \sum_0^{n-1} \chi_{K^c}(X(m)))] \\ &\leq e^n + e^{nM} Q_{n,x}(C(L)^c) . \end{aligned}$$

Since $Q_{n,x}(C(L)^c) \leq e^{-nM}$ for large n's , it is clear that

$$\overline{\lim_{n\to\infty}} \frac{1}{n} \log E^{P_x}[\exp(M \sum_0^{n-1} \chi_{K^c}(X(m)))] \leq 1 .$$

Since we can do this for every $x \in E$, we now see that if $V \in C_b(E)$ vanishes on $K(M)$ and if $\|V\|_{C_b(E)} \leq M$, then $\tilde{\lambda}(V) \leq 1$. $\square$

(8.15) Exercise: Let $\Phi : E \to R^1$ be given. In case $(\underline{C})$, show that (8.10) holds if there exists a sequence $\{u_n\}_1^\infty \subseteq D(A)$ such that: $u_n \geq \varepsilon$ for some $\varepsilon > 0$ and all n, $\sup_{n\geq 1} \sup_{x\in\Gamma} u_n(x) < \infty$, and $-\frac{Au_n}{u_n} \to \Phi$ point-wise. (What is the analogous criterion in case $(\underline{D})$?) Next, suppose that $E = R^d$ and that $D(A)$ contains the space $C_0^2(R^d)$ of $\phi \in C^2(R^d)$ having compact support. Further, assume that for $\phi \in C_0^2(R^d)$, $A\phi = L\phi$, where L is a (degenerate) elliptic operator

$$1/2 \sum_{i,j=1}^{d} a^{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} + \sum_{i=1}^{d} b^i(x)\frac{\partial}{\partial x_i}$$

having continuous coefficients. Show that (8.10) holds for every $\Gamma \subset\subset R^d$ as soon as there exists a uniformly positive $u \in C^2(R^d)$ such that $-\frac{Lu}{u} = \Phi$.

From now on, we are going to restrict our attention to the case $(\underline{S}.\underline{C}.)$.

(8.16) Lemma: Assume $(\underline{S}.\underline{C}.)$ with $m \in \mathcal{M}_1(E)$. Also, assume that J_σ is a rate function (i.e. $\{J_\sigma \leq L\} \subset\subset \mathcal{M}_1(E)$ for all $L > 0$). Then λ_σ is tight, $\lambda_\sigma = J_\sigma^*$, $J_\sigma = \lambda_\sigma^*$, and for all closed $F \subseteq \mathcal{M}_1(E)$:

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log Q_{t,m}(F) \leq -\inf_{\nu\in F} J_\sigma(\nu) \quad , \tag{8.17}$$

where $Q_{t,m} = \int Q_{t,x} m(dx)$. Next, add the assumption that $P(t,x,\cdot) \ll m$ for all $t > 0$ and $x \in E$. Then $J = J_\sigma$, and so the preceding statements hold with J replacing J_σ. Finally, add the assumption that for each $x \in E$ there is a $T > 0$ such that $p(T,x,\cdot) \equiv \frac{P(T,x,dy)}{m(dy)} \in L^q(m)$ for some $q \in (1,\infty]$ which is independent of x. Then, $\tilde{\lambda}$ is tight and $\tilde{\lambda} = \lambda_\sigma$. In particular, Theorem (8.6) applies in this situation.

Proof: By (7.43), we always have that $\lambda_\sigma = J_\sigma^*$. Thus, if J_σ is a rate function (and therefore is l.s.c.) $J_\sigma = \lambda_\sigma^*$ follows from Theorem (7.15). Also, if J_σ is a rate function, then, from $\lambda_\sigma = J_\sigma^*$ and the first part of Lemma (8.5), we will know that λ_σ is tight as soon as we show that $J_\sigma(\nu_0) = 0$ for some $\nu_0 \in \mathcal{M}_1(E)$. But $1 \in L^2(m)$ and $P_t 1 = 1$, $t > 0$. Thus $\varepsilon(1^{1/2}, 1^{1/2}) = 0$, and so $J_\sigma(m) = 0$. Once one knows that λ_σ is tight, the proof of (8.17) is precisely the same as that of Theorem (8.6).

Next, suppose that $P(t,x,\cdot) \ll m$ for $t > 0$ and $x \in E$. Then, by Theorem (7.44), $J = J_\sigma$.

Finally, to prove the last assertion, we need only check that $\tilde{\lambda}$ is tight under the stated hypothesis. Indeed, if $\tilde{\lambda}$ is tight, then by Theorem (8.6), $\tilde{\lambda} = J^*$ and so $\tilde{\lambda} = J^* = J_\sigma^* = \lambda_\sigma$. Since λ_σ is tight, to prove that $\tilde{\lambda}$ is tight, it suffices to show that $\lambda(V) \le 1/\alpha\, \lambda_\sigma(\alpha V)$ for all $V \in C_b(E)$ and some $\alpha \in (1,\infty)$. To this end, let $x \in E$ be given and choose $T > 0$ so that $p(T,x,\cdot) \in L^q(m)$. Then, for any $V \in C_b(E)$:

$$\overline{\lim_{t\to\infty}} \frac{1}{t} \log E^{P_x}[\exp(\int_0^t V(X(s))ds)]$$

$$= \overline{\lim_{t\to\infty}} \frac{1}{t} \log E^{P_x}[\exp(\int_T^{T+t} V(X(s))ds)]$$

$$= \overline{\lim_{t\to\infty}} \frac{1}{t} \log(\int p(T,x,y) E^{P_y}[\exp(\int_0^t V(X(s))ds)] m(dy))$$

$$\le \overline{\lim_{t\to\infty}} \frac{1}{t} \log[\|p(T,x,\cdot)\|_{L^q(m)} (\int (E^{P_y}[\exp(\int_0^t V(X(s))ds)])^{q'} m(dy))^{1/q'}]$$

$$\leq \frac{1}{q'} \overline{\lim_{t\to\infty}} \frac{1}{t} \log(\int E^{P_y}[\exp(q'\int_0^t V(X(s))ds)]m(dy))$$

$$= \frac{1}{q'} \overline{\lim_{t\to\infty}} \frac{1}{t} \log(1, \bar{P}_t^{q'V} 1)_{L^2(m)} \leq 1/q' \ \lambda_\sigma(q'V) \quad ,$$

where $1/q' = 1 - 1/q$. Thus we can take $\alpha = q'$.

□

We now turn to the problem of proving a lower bound. The approach which we are going to take works only in the case (S.C.). We begin with a simple version of the general Cameron-Martin transformation.

(8.18) Lemma: In case (C) , suppose that $u \in \mathcal{U} \cap D(A)$ and set $V_u = -\frac{Au}{u}$. Define

$$R^u(t) = \frac{u(X(t))}{u(X(0))} \exp(\int_0^t V_u(X(s))ds) \quad .$$

Then, for each $x \in E$, $(R^u(t), \mathcal{M}_t, P_x)$ is a martingale (recall that $\mathcal{M}_t \equiv \sigma(X(s) : 0 \leq s \leq t)$). In particular, for each $x \in E$, there is a unique $Q_x^u \in \mathcal{M}_1(\Omega)$ such that $Q_x^u(B) = E^{P_x}[R^u(t), B]$ for all $t \geq 0$ and $B \in \mathcal{M}_t$. Moreover, $\{Q_x^u : x \in E\}$ is a Feller continuous, time homogeneous Markov family with associated semigroup $\{Q_t^u : t > 0\}$ given by $[Q_t^u \phi](x) = \frac{1}{u(x)}[P_t^{V_u}(u \cdot \phi)](x)$. Finally, if $\{P_t : t > 0\}$ satisfies (S.C.) with respect to $m \in \mathcal{M}_1(E)$, then $\{Q_t^u : t > 0\}$ satisfies (S.C.) with respect to the measure m_u given by $m_u(dy) = u(y)^2 m(dy)$.

Proof: First note that $\frac{d}{dt}[P_t^{V_u} u] = 0$ and, therefore, that $Q_t^u 1 = 1$. Next, given $B \in \mathcal{M}_{t_1}$:

$$E^{P_x}[R^u(t_1+t_2)\phi(X(t_1+t_2)), B]$$

$$= E^{P_x}[R^u(t_1)(R^u(t_2)\phi(X(t_2)))\circ\theta_{t_1},B]$$

$$= E^{P_x}[R^u(t_1)E^{P_{x(t_1)}}[R^u(t_2)\phi(X(t_2))],B]$$

for all $\phi \in C_b(E)$. In particular, when $\phi = 1$, we conclude that $(R^u(t),\mathcal{M}_t,P_x)$ is a martingale. The existence of Q_x now follows from standard extension theorems. Next, note that the preceding can be re-written as:

$$E^{Q_x^u}[\phi(X(t_1+t_2)),B]$$

$$= E^{Q_x^u}[[Q_{t_2}^u\phi](X(t_1)),B] \quad .$$

Since $x \to Q_x^u$ is continuous, this proves that $\{Q_x^u : x \in E\}$ is a Feller continuous, time homogeneous Markov family with associated semigroup $\{Q_t^u : t > 0\}$.

Finally, suppose that $\{P_t : t > 0\}$ satisfies (S.C) with respect to $m \in \mathcal{M}_1(E)$. Then,

$$\int\phi_1\cdot\overset{V}{P_t}{}^u\phi_2\,dm = \int\phi_2\cdot\overset{V}{P_t}{}^u\phi_1\,dm$$

for all $t > 0$ and $\phi_1,\phi_2 \in C_b(E)$ (cf. Lemma (7.29)). Hence

$$\int\phi_1\cdot Q_t^u\phi_2\,dm_u = \int u\phi_1\cdot\overset{V}{P_t}{}^u(u\phi_2)dm$$

$$= \int u\phi_2\cdot\overset{V}{P_t}{}^u(u\phi_1)dm = \int\phi_2\cdot Q_t^u\phi_1\,dm_u \quad . \qquad \square$$

(8.19) Lemma: In case (S.C.) with $m \in \mathcal{M}_1(E)$, suppose that the only $\phi \in L^2(m)$ satisfying $\varepsilon(\phi,\phi) = 0$ are m-almost surely constant. Given

$u \in \mathcal{U} \cap D(A)$ with $\int u^2 dm = 1$, define $\{Q_x^u : x \in E\}$ and m_u as in Lemma (8.18) and set $Q^u = \int Q_x^u m_u(dx)$. Then $Q^u = Q^u \circ \theta_t^{-1}$ for all $t > 0$ and Q^u is $\theta_\bullet$-ergodic (i.e. for any $B \in \mathcal{M}$ satisfying $B = \theta_t^{-1}(B)$ (a.s., Q^u) for all $t > 0$, $Q^u(B) \in \{0,1\}$). In particular, for any open $G \ni m_u$, $Q^u(L_t \in G) \to 1$ as $t \to \infty$.

<u>Proof</u>: Clearly $Q^u = Q^u \circ \theta_t^{-1}$, $t > 0$, since m_u is $\{Q_t^u : t > 0\}$ stationary ($\int Q_t^u \phi dm_u = \int (Q_t^u 1) \cdot \phi dm_u = \int \phi dm_u$). We next show that if $\phi \in L^2(m_u)$ and $\overline{Q_t^u \phi} = \phi$ (a.s., m_u) for $t > 0$, then $\phi = \int \phi dm_u$ (a.s., m_u). Indeed, $\phi \in L^2(m)$ and (cf. Lemma (7.38)):

$$\frac{1}{t}\,(\phi, \phi - \overline{P}_t \phi)_{L^2(m)} = \frac{1}{2} \int (\phi(y) - \phi(x))^2 m_t(dx \times dy)$$

where $m_t(dx \times dy) = P(t,x,dy)m(dx)$. Also, if $Q^u(t,x,\cdot)$ is the transition function for $\{Q_x^u : x \in E\}$ and $m_t^u(dx \times dy) = Q^u(t,x,dy)m_u(dx)$, then

$$0 = \frac{1}{t}\,(\phi - \overline{Q_t^u \phi}, \phi)_{L^2(m_u)} = \frac{1}{2} \int (\phi(y) - \phi(x))^2 m_t^u(dx \times dy) \quad .$$

Note that

$$m_t^u(\Gamma_1 \times \Gamma_2) = \int_{\Gamma_1} Q^u(t,x,\Gamma_2) m_u(dx)$$

$$= \int_{\Gamma_1} u(x) [P_t^{V_u}(u \cdot \chi_{\Gamma_2})](x) m(dx)$$

$$\geq \varepsilon^2 \exp\left(-t \|V_u\|_{C_b(E)}\right) \int_{\Gamma_1} P_t(t,x,\Gamma_2) m(dx)$$

$$= \varepsilon^2 \exp\left(-t \|V_u\|_{C_b(E)}\right) m_t(\Gamma_1 \times \Gamma_2) \quad ,$$

where $0 < \varepsilon \leq u$. Thus:

$$\lim_{t\to 0} \frac{1}{t}(\phi,\phi-\bar{P}_t\phi)_{L^2(m)} \leq \frac{1}{\varepsilon^2}\exp\left(t\|V_n\|_{C_b(E)}\right)\overline{(\phi,\phi-Q_t^u\phi)}_{L^2(m_u)} = 0 ,$$

and so (cf. Lemma (7.38)), $\varepsilon(\phi,\phi) = 0$. By hypothesis, this means that $\phi = \text{const.}$ (a.s., m) and so $\phi = \int\phi dm_u$ (a.s., m_u).

To prove that Q^u is $\theta_\bullet$-ergodic, we must show that if $\Phi : \Omega \to R^1$ is a bounded, $\mathcal{M}$-measurable function satisfying $\Phi = \Phi\circ\theta_t$ (a.s., Q^u) for all $t > 0$, then $\Phi = E^{Q^u}[\Phi]$ (a.s., Q^u). To this end, set $\phi(x) = E^{Q_x^u}[\Phi]$, $x \in E$. Then $Q_t^u\phi(x) = E^{Q_x^u}[\phi(X(t))] = E^{Q_x^u}[E^{Q_{X(t)}^u}[\Phi]] = E^{Q_x^u}[\Phi\circ\theta_t] = E^{Q_x^u}[\Phi] = \phi(x)$ for m_u-almost all $x \in E$. Thus, $\phi = \int\phi dm_u$ (a.s., m_u); and so, for each $t > 0$, $\phi(X(t)) = \int\phi dm_u = E^{Q^u}[\Phi]$ (a.s., Q^u). Hence, if $t > 0$ and $B \in \mathcal{M}_t$, then $E^{Q^u}[\Phi,B] = E^{Q^u}[\Phi\circ\theta_t,B] = E^{Q^u}[\phi(X(t)),B] = E^{Q^u}[\Phi]Q^u(B)$. Since this is true for all $t > 0$ and $B \in \mathcal{M}_t$, this proves that $\Phi = E^{Q^u}[\Phi]$ (a.s., Q^u).

Finally, we now know, by the individual ergodic theorem, that $\frac{1}{t}\int_0^t \phi(X(s))ds \to \int\phi dm_u$ (a.s., Q^u) for each bounded measurable $\phi : E \to R^1$. Clearly, this implies that $Q^u(L_t \in G) \to 1$ whenever G is an open set containing m_u (cf. our introduction discussion in section 5) . □

(8.20) Lemma: In case (S.C.) with $m \in \mathcal{M}_1(E)$, let $f \in L^1(m)^+$ with $\int f dm = 1$ and $\varepsilon(f^{1/2},f^{1/2}) < \infty$ be given. Then there exists a sequence $\{u_n\}_1^\infty \subseteq \mathcal{U} \cap D(A)$ such that $\int u_n^2 dm = 1$, $u_n \to f^{1/2}$ in $L^2(m)$, and $\varepsilon(u_n,u_n) \to \varepsilon(f^{1/2},f^{1/2})$.

Proof: Clearly it is enough to prove the result without requiring that $\int u_n^2 dm = 1$, since we can always replace u_n by $u_n/\|u_n\|_{L^2(M)}$ if necessary.

Thus we will ignore the normalization.

First, note that $\overline{P}_t f^{1/2} \in D(\overline{A})^+$ for all $t > 0$ and that $\overline{P}_t f^{1/2} \to f^{1/2}$ in $L^2(m)$ and $\varepsilon(\overline{P}_t f^{1/2}, \overline{P}_t f^{1/2}) \to \varepsilon(f^{1/2}, f^{1/2})$ as $t \downarrow 0$. Thus we need only prove the result when $f^{1/2} \in D(\overline{A})^+$.

For $\lambda > 0$, define $R_\lambda = \int_0^\infty e^{-\lambda t} P_t dt$ and $\overline{R}_\lambda = \int_0^\infty e^{-\lambda t} \overline{P}_t dt$, respectively. Clearly $\overline{R}_\lambda$ is the unique continuous extension of R_λ to $L^2(m)$. Given $f^{1/2} \in D(\overline{A})^+$, there is a $\psi \in L^2(m)$ such that $f^{1/2} = \overline{R}_1 \psi$. Now choose $\{\psi_n\}_1^\infty \subseteq C_b(E)$ so that $\psi_n \to \psi$ in $L^2(m)$, and set $\phi_n = R_1 \psi_n$. Then $\phi_n \in D(A)$, $\phi_n \to f^{1/2}$ in $L^2(m)$, and $A\phi_n \to \overline{A} f^{1/2}$ in $L^2(m)$. In particular, $\varepsilon(\phi_n, \phi_n) \to \varepsilon(f^{1/2}, f^{1/2})$. Moreover, since $f^{1/2} \in L^2(m)^+$, $\tilde{\phi}_n \equiv \phi_n \vee 0 + 1/n \to f^{1/2}$ in $L^2(m)$ and therefore $\varepsilon(f^{1/2}, f^{1/2}) \leq \varliminf_{n \to \infty} \varepsilon(\tilde{\phi}_n, \tilde{\phi}_n)$. At the same time, from Lemma (7.38), it is clear that $\varepsilon(\tilde{\phi}_n, \tilde{\phi}_n) \leq \varepsilon(\phi_n, \phi_n)$ for all $n \geq 1$. Hence $\varepsilon(f^{1/2}, f^{1/2}) = \lim_{n \to \infty} \varepsilon(\tilde{\phi}_n, \tilde{\phi}_n)$. We therefore see that it suffices to handle the case in which $f^{1/2} \in \mathcal{U}$ and $\varepsilon(f^{1/2}, f^{1/2}) < \infty$. But in this case, we take $u_n = nR_n f^{1/2}$. Clearly $u_n \in \mathcal{U} \cap D(A)$; and, from the spectral theorem, we see that $u_n \to f^{1/2}$ in $L^2(m)$ while $\varepsilon(u_n, u_n) \to \varepsilon(f^{1/2}, f^{1/2})$. □

(8.21) <u>Theorem</u>: In case (<u>S</u>.<u>C</u>.) with $m \in \mathcal{M}_1(E)$, suppose that the only $\phi \in L^2(m)$ with $\varepsilon(\phi, \phi) = 0$ are m-almost surely constant. Define $Q_{t,m} = \int Q_{t,x} m(dx)$. Then, for all open G in $\mathcal{M}_1(E)$,

$$\varliminf_{t \to \infty} \frac{1}{t} \log Q_{t,m}(G) \geq -\inf_{\nu \in G} J_\sigma(\nu) . \tag{8.22}$$

Moreover, if, in addition, $\Gamma \subseteq E$ is a set such that for some $T > 0$ and all

$\varepsilon > 0$ there exists a $\delta > 0$ for which $m(\Delta) < \varepsilon$ whenever $\inf_{x\in\Gamma} P(T,x,\Delta) < \delta$, then

(8.23) $$\underline{\lim_{t\to\infty}} \frac{1}{t} \log(\inf_{x\in\Gamma} Q_{t,x}(G)) \geq -\inf_{\nu\in G} J_\sigma(\nu) \quad .$$

<u>Proof</u>: By Lemma (8.20), (8.22) and (8.23) will follow, respectively, once we show that for each $u \in \mathfrak{U} \cap D(A)$ satisfying $m_u \in G$:

(8.24) $$\lim_{t\to\infty} \frac{1}{t} \log Q_{t,m}(G) \geq -\varepsilon(u,u)$$

and

(8.25) $$\underline{\lim_{t\to\infty}} \frac{1}{t} \log (\inf_{x\in\Gamma} Q_{t,x}(G)) \geq -\varepsilon(u,u) \quad .$$

Let $\varepsilon > 0$ be chosen so that $\varepsilon \leq u \leq 1/\varepsilon$. For any open B, we have:

$$\begin{aligned} Q_{t,m}(B) &\geq \varepsilon^2 \int Q_{t,x}(B) m_u(dx) \\ &\geq \varepsilon^2 E^{P_{m_u}}[\exp(\int_0^t V_u(X(s))ds), L_t \in B] \\ &\quad \times \inf_{\nu\in B} \exp(-t\int V_u d\nu) \\ &\geq \varepsilon^4 Q^u(L_t \in B)\exp(-\sup_{\nu\in B} t\int V_u d\nu) \quad . \end{aligned}$$

Thus, if $m_u \in B$, then, since $Q^u(L_t \in B) \to 1$,

$$\underline{\lim_{t\to\infty}} \frac{1}{t} \log Q_{t,m}(B) \geq -\sup_{\nu\in B} \int V_u d\nu \quad .$$

Choosing $\{B_n\}_1^\infty$ so that $G \supseteq B_n \downarrow \{m_u\}$, we conclude that

$$\varliminf_{t\to\infty} \frac{1}{t} \log Q_{t,m}(G) \geq \int \frac{Au}{u}\, dm_u = \int u\cdot Au dm = -\varepsilon(u,u) \quad .$$

Thus, (8.24) has been proved.

The proof of (8.25) is similar. Namely, for any open $B \subsetneq G$ which is a positive distance from G^c, $Q_{t,x}(G) \geq P_x(L_t \circ \theta_T \in B)$ for all $x \in E$ and all sufficiently large t's. Moreover,

$$\begin{aligned} P_x(L_t \circ \theta_T \in B) &= \int P_y(L_t \in B) P(T,x,dy) \\ &\geq \varepsilon^2 \int Q_y^u(L_t \in B) P(T,x,dy) \\ &\quad \times \exp(-\sup_{\nu\in B} t\int V_u d\nu) \quad . \end{aligned}$$

Note that since $Q^u(L_t \in B) = \int Q_y^u(L_t \in B) m_u(dy) \to 1$ if $m_u \in B$, $Q_\cdot^u(L_t \in B) \to 1$ in m-measure for such B's. Thus, our hypothesis implies that there is a $\delta > 0$ such that $\inf_{x\in\Gamma} P(T,x,\{y : Q_y^u(L_t \in B) \geq 1/2\}) \geq \delta$ for all sufficiently large t's. Combining these remarks, we see that if $B \subseteq G$ is open, $\operatorname{dist}(B,G^c) > 0$, and $m_u \in B$, then

$$\varliminf_{t\to\infty} \frac{1}{t} \log(\inf_{x\in\Gamma} Q_{t,x}(G)) \geq -\sup_{\nu\in B} \int V_u d\nu \quad .$$

From here, the argument is exactly the same as for (8.24).

□

(8.26) <u>Corollary</u>: In case (<u>S.C.</u>) with $m \in \mathcal{M}_1(E)$, assume that J_σ is a rate function and that the only $\phi \in L^2(m)$ with $\varepsilon(\phi,\phi) = 0$ are m-almost surely constant. Then, $\{Q_{t,m} : t > 0\}$ satisfies the large deviation principle with rate function J_σ. Next, add the assumption that, for all $t > 0$ and $x \in E$, $P(t,x,\cdot) << m$ and that, for each $x \in E$ and some $T > 0$, $m << P(T,x,\cdot)$. Then, $J = J_\sigma$ and, for each $x \in E$, $\{Q_{t,x} : t > 0\}$

satisfies the large deviation principle with rate J. In particular, $\tilde{\lambda} = \lambda_\sigma = J^*$ and $\tilde{\lambda}$ is tight. Finally, let $\Gamma \subseteq E$ and add the assumptions that (8.8) holds for some $\Phi : E \to R^1$ which is bounded below and satisfies $\{\Phi \leq L\} \subset\subset E$ for all $L > 0$ and that for some $T > 0$ and all $\varepsilon > 0$ there is a $\delta > 0$ such that $m(\Delta) < \varepsilon$ whenever $\inf\limits_{x \in \Gamma} P(T,x,\Delta) < \delta$. Then $\{Q_{t,x} : t > 0\}$ satisfies the large deviation principle with rate J uniformly for $x \in \Gamma$ (i.e. (8.14) holds for each closed F and (8.23) holds for each open G).

Proof: Since all these assertions are simply re-statements of results already derived, there is nothing to prove here.

$\square$

(8.27) Exercise: Let $\{P_x : x \in E\}$ with associated transition probability function $P(t,x,\cdot)$ satisfy (S.C.) with respect to $m \in \mathcal{M}_1(E)$. Assume that $P(t,x,dy) = p(t,x,y)m(dy)$, $t > 0$ and $x \in E$, where $p(t,\cdot,\cdot) \in C(E \times E)$. Show that $p(t,x,y) = p(t,y,x)$ and therefore that $p(2t,x,x) = \int |p(t,x,y)|^2 m(dy)$. Next, show that $J = J_\sigma$ is a rate function if $\int p(t,x,x)m(dx) < \infty$ for all $t > 0$. (Hint: check that $\bar{P}_t : L^2(m) \to L^2(m)$ is compact for each $t > 0$ and show that

$$\lim_{t \downarrow 0} \sup_{\substack{\phi \in L^2(m) \\ \varepsilon(\phi,\phi) \leq L}} \|\bar{P}_t \phi - \phi\|_{L^2(m)} = 0$$

for all $L > 0$.) Next, let $\Gamma \subseteq E$ be given. Show that for all $\varepsilon > 0$ there exists a $\delta > 0$ such that $m(\Delta) < \varepsilon$ whenever $\inf\limits_{x \in \Gamma} P(T,x,\Delta) < \delta$ if $p(T,x,\cdot) > 0$ (a.s., m) for each $x \in \Gamma$ and either one of the following conditions holds:

i) there is an $\alpha > 0$ such that

$$\sup_{x\in\Gamma} \int \left|p(T,x,y)\right|^{-\alpha} m(dy) < \infty \quad ,$$

ii) as a map from Γ into $L^1(m)$, $x \to p(T,x,\cdot)$ is uniformly continuous.

(8.28) Exercise: Let $E = R^1$. For each $x \in R^1$, denote by $X(\cdot)$ the solution to the linear equation:

$$X(t,x) = x + \beta(t) - 1/2 \int_0^t X(s,x)ds \quad , \quad t \geq 0 \quad ,$$

where $\beta(\cdot)$ is a 1-dimensional Brownian motion starting at 0 . Show that:

$$X(t,x) = xe^{-t/2} + \int_0^t e^{-(t-s)/2} d\beta(s) \quad ,$$

and thereby conclude that $X(\cdot,x)$ is a Gaussian process. Next, let P_x on $C([0,\infty);R^1)$ denote the distribution of $X(\cdot,x)$. Show that $\{P_x : x \in R^1\}$ is a Feller continuous, time-homogeneous Markov family with transition function $P(t,x,\cdot)$ given by:

$$P(t,x,dy) = \frac{1}{(1-e^{-t})^{1/2}} \exp(-(y-xe^{-t/2})^2/2(1-e^{-t}))dy \quad .$$

Also, show that $\{P_x : x \in R^1\}$ satisfies (S.C.) with respect to $\gamma(dy) = (2\pi)^{-1/2}e^{-y^2/2}dy$.

Now show that if A is the weak generator of the associated semigroup $\{P_t : t > 0\}$ on $C_b(R^1)$, then $C_0^2(R^1) \subseteq D(A)$ and $A\phi = L\phi$ for $\phi \in C_0^2(R^1)$ where $L = 1/2(\frac{\partial^2}{\partial x^2} - x\frac{\partial}{\partial x})$. Moreover, if $u(x) = e^{x^2/4}$, check that

$$-\frac{Lu}{u} = \frac{1}{4}(1/2\, x^2 - 1) \quad .$$

We conclude, using (8.15), that

$$(8.29) \qquad \sup_{t>0} E^{P_x}[\exp(\tfrac{1}{4}\int_0^t (1/2X(s)^2 - 1)ds)] \leq e^{x^2/4}, \quad x \in R^1 .$$

Finally using (8.27) and the above, conclude that for each $K \subset\subset R^1$, $\{Q_{t,x} : t > 0\}$ satisfies the large deviation principle uniformly for $x \in K$ with rate function I described as follows: $I(\nu) = \infty$ unless $\nu \in \mathcal{M}_1(E)$ and $\nu(dy) = f(y)\gamma(dy)$ where $f^{1/2}$ has one distibutional derivative $(f^{1/2})' \in L^2(\gamma)$, in which case $I(\nu) = 1/2\|(f^{1/2})'\|^2_{L^2(\gamma)}$.

9. Logarithmic Sobolev Inequalities:

There is an interesting connection between our considerations here and L. Gross's theory of logarithmic Sobolev inequalities. For our purposes, it is best to describe a logarithmic Sobolev inequality in the following terms. Let $\{P_x : x \in E\}$ satisfy (S.C.) with respect to $m \in \mathcal{M}_1(E)$. A logarithmic Sobolev inequality is a statement of the form:

$$(9.1) \qquad J_m \leq \alpha J_\sigma$$

for some $\alpha > 0$, where $J_m : \mathcal{M}_1(E) \to [0,\infty) \cup \{\infty\}$ is defined by:

$$J_m(\nu) = \begin{cases} \int \log(\frac{d\nu}{dm})d\nu & \text{if } \nu \in \mathcal{M}_1(E) \text{ and } \nu << m \\ \infty & \text{otherwise.} \end{cases}$$

Obviously, (9.1) has interesting implications for the large deviation theory associated with $\{P_x : x \in E\}$.

To begin with, we note that (9.1) implies that the set

$$(9.2) \qquad \mathcal{F}_L = \{f \in L^1(m)^+ : \|f\|_{L^1(m)} = 1 \text{ and } \varepsilon(f^{1/2},f^{1/2}) \leq L\}$$

is a convex, weakly compact subset of $L^1(m)$. Indeed by Lemma (7.40), $\mathcal{F}_L$ is closed and convex. Moreover, (9.1) implies that $\mathcal{F}_L \subseteq \{f \in L^1(m)^+ : \int f \log f dm \leq \alpha L\}$. Since $\xi \log \xi \geq -e^{-1}$, $\xi \geq 0$, it follows that $\mathcal{F}$ is a uniformly m-integrable subset of $L^1(m)$. Since it is closed and convex, we now see that $\mathcal{F}_L$ is weakly compact. We next show that (9.1) implies that

$$(9.3) \qquad \{\nu \in \mathcal{M}(E) : J_\sigma(\nu) \leq L\} \subset\subset \mathcal{M}_1(E) , \quad L > 0 .$$

Indeed, let $\{\nu_n\}_1^\infty \subseteq \mathcal{M}(E)$ satisfy $\sup_n J_\sigma(\nu_n) \leq L$. Choose $\{f_n\}_1^\infty \subseteq \mathcal{F}_L$ so that $\nu_n(dx) = f_n(x)m(dx)$. Then there is a subsequence $\{f_{n'}\}$ such that $f_{n'} \to f$ weakly in $L^1(m)$. Clearly $\nu_{n'} \Rightarrow \nu$ where $\nu(dx) = f(x)m(dx)$. To see that $J_\sigma(\nu) \leq L$, simply note that, since $\mathcal{F}_L$ is closed and convex $f \in \mathcal{F}_L$. As a consequence of (9.3) and the first part of Lemma (8.16), we now see that (9.1) implies that:

$$(9.4) \qquad \overline{\lim_{t\to\infty}} \frac{1}{t} \log Q_{t,m}(F) \leq -\inf_{\nu\in F} J_\sigma(\nu) \leq -\frac{1}{\alpha} \inf_{\nu\in F} J_m(\nu)$$

for all closed $F \subseteq \mathcal{M}_1(E)$.

We next observe that (9.1) is equivalent to the statement that

$$(9.5) \qquad \lambda_\sigma(V) \leq \frac{1}{\alpha} \lambda_m(\alpha V) , \quad V \in C_b(E) ,$$

where

$$\lambda_m(V) \equiv \log\left(\int e^{V(y)} m(dy)\right) .$$

To see that (9.5) follows from (9.1), note that, by Lemma (3.37), $J_m = \lambda_m^*$ and that λ_m satisfies the hypotheses of Lemma (8.2). Thus, $\lambda_m = J_m^*$. At the same time, by (7.43), $\lambda_\sigma = J_\sigma^*$. Thus, from (9.1), we

deduce that:

$$\frac{1}{\alpha}\lambda_m(\alpha V) = \sup_\nu(\int V d\nu - \frac{1}{\alpha}J_m(\nu))$$

$$\geq \sup_\nu(\int V d\nu - J_\sigma(\nu)) = \lambda_\sigma(V) \quad ;$$

and so, (9.1) implies (9.5) . Conversely, suppose that (9.5) holds. Given $u \in \mathcal{U} \cap D(A)$ satisfying $\|u\|_{L^2(m)} = 1$, (9.5) implies that

$$(u,A_V u)_{L^2(m)} \leq \frac{1}{\alpha}\lambda_m(\alpha V)$$

for all $V \in C_b(E)$. In particular, if $V = \frac{1}{\alpha}\log u^2$, then we get:

$$-\varepsilon(u,u) + \frac{1}{\alpha}\int u^2 \log u^2 dm \leq 0 \quad ,$$

since $\lambda_m(\log u^2) = \log\|u\|_2 = 0$. In other words,

$$\int u^2 \log u^2 dm \leq \alpha\varepsilon(u,u)$$

for all such u's . Applying Lemma (8.20) , it is now clear that $\int \phi\log\phi dm \leq \alpha\varepsilon(\phi^{1/2},\phi^{1/2})$ for all $\phi \in L^1(m)^+$ satisfying $\|\phi\|_{L^1(m)} = 1$; and this is precisely (9.1) . That (9.1) is equivalent to (9.5) was already noticed by L. Gross in a slightly different context.

We next use the equivalence between (9.1) and (9.5) to give an initial link between logarithm Sobolev inequalities and hypercontraction results. A hypercontraction result is a statement of the form:

$$\text{(9.6)} \qquad \|P_T\phi\|_{L^q(m)} \leq \|\phi\|_{L^p(m)} \quad , \quad \phi \in C_b(E) \quad ,$$

for some $T > 0$, $p \in (1,\infty)$, and $q \in (p,\infty)$. What we are now going to do is show that (9.6) implies (9.5) and therefore (9.1), with α given by:

(9.7) $$\alpha(T,p,q) = \frac{pqT}{q-p} \quad .$$

To see this, first note that

$$\lambda_\sigma(V) = \lim_{t\to\infty} \frac{1}{t} \log(\int [P_t^V 1](x)m(dx)) \quad .$$

Indeed, it is clear that the right side of the above is dominated by the left. On the other hand, if $\lambda < \lambda_\sigma(V)$, then we can find a $\phi \in C_b(E)$ such that $\lim_{t\to\infty} \frac{1}{t} \log(\phi,P_t^V\phi) \geq \lambda$. Since $\|\phi\|^2_{C_b(E)} \int[P_t^V 1](x)m(dx) \geq (\phi,P_t\phi)_{L^2(m)}$, we now have the desired equality. Next, note that, just as in the proof of Corollary (7.24) :

$$\begin{aligned}
\int[P_{nT}^{V/T} 1](x)m(dx) &= E^{P_m}[\exp(\frac{1}{T}\int_0^{nT} V(X(s))ds)] \\
&\leq \int_0^1 E^{P_m}[E^{P_{x(tT)}}[\exp(\sum_{k=0}^{n-1} V(X(kT)))]]dt \\
&= E^{P_m}[\exp(\sum_{k=0}^{n-1} V(X(kT)))] \\
&= \int[((P_T)_V)^n 1](x)m(dx) \quad ,
\end{aligned}$$

where $(P_T)_V = e^V P_T$. Combining this with the preceding, we arrive at:

$$\lambda_\sigma(V/T) \leq \frac{1}{T} \varliminf_{n\to\infty} \frac{1}{n} \log\|((P_T)_V)^n 1\|_{L^1(m)}$$

In particular,

(9.8) $$\lambda_\sigma(V/T) \leq \frac{1}{T} \log\|(P_T)_V\|_{Hom(L^p(m);L^p(m))} \quad .$$

But, by Hölder's inequality:

$$\|(P_T)_V\psi\|_{L^p(m)} = \left(\int e^{pV(x)}[P_T\psi]^p(x)m(dx)\right)^{1/p}$$

$$\leq \left(\int \exp\left(\frac{pq}{q-p}V(x)\right)m(dx)\right)^{\frac{q-p}{pq}} \|P_T\psi\|_{L^q(m)} ,$$

and so (9.6) implies that

$$\log\|(P_T)_V\|_{Hom(L^p(m);L^p(m))} < \frac{q-p}{pq}\lambda_m\left(\frac{pq}{q-p}V\right) .$$

Combining this with (9.8) , we now see that (9.6) implies (9.5) with α given by (9.7) .

We next turn to Gross's theorem which shows that hypercontraction results from a logarithmic Sobolev inequality.

(9.9) <u>Lemma</u>: For any $p \in (1,\infty)$ and $u \in \mathfrak{U} \cap D(A)$,
$\varepsilon(u^{p/2},u^{p/2}) \leq \frac{p^2}{4(p-1)}(u^{p-1},-Au)_{L^2(m)}$.

<u>Proof</u>: Given $t > 0$ and $\phi,\psi \in L^2(m)$, define
$\varepsilon_t(\phi,\psi) = \frac{1}{t}(\phi - \overline{P}_t\phi,\psi)_{L^2(m)}$. Then, by Lemma (7.38) , for $\phi,\psi \in D(\overline{A})$:

$$(-\overline{A}\phi,\psi)_{L^2(m)} = \frac{1}{4}(\varepsilon(\phi+\psi,\phi+\psi) - \varepsilon(\phi-\psi,\phi-\psi))$$

$$= \lim_{t\downarrow 0} \varepsilon_t(\phi,\psi) .$$

Note that for any $\phi \in D(\overline{A})$ and $\psi \in L^2(m)$:

$$\left|\varepsilon_t(\phi,\psi)\right| = \left|\frac{1}{t}(\phi-\overline{P}_t\phi,\psi)_{L^2(m)}\right|$$

$$\leq \frac{1}{t}\|\phi-\overline{P}_t\phi\|_{L^2(m)}\|\psi\|_{L^2(m)} \leq \|\overline{A}\phi\|_{L^2(m)}\|\psi\|_{L^2(m)} .$$

Thus, for any $\phi \in D(\overline{A})$ and $\psi \in L^2(m)$:

$$(-\overline{A}\phi,\psi)_{L^2(m)} = \lim_{t\downarrow 0} \varepsilon_t(\phi,\psi)$$

$$= \frac{1}{2} \lim_{t\downarrow 0} \frac{1}{t} \int(\phi(y)-\phi(x))(\psi(y)-\psi(x))m_t(dx \times dy) \quad ,$$

where $m_t(dx \times dy) = P(t,x,dy)m(dx)$ and we have used the results of Lemma (7.38) (note that $\sigma_t \equiv 1$ since $m \in \mathcal{M}_1(E)$). At the same time:

$$\varepsilon(\phi,\phi) = \frac{1}{2} \lim_{t\downarrow 0} \frac{1}{t} \int(\phi(u)-\phi(x))^2 m_t(dx \times dy)$$

for all $\phi \in L^2(m)$. Thus, we need only show that

$$\int(u^{p/2}(y)-u^{p/2}(x))^2 m_t(dx \times dy)$$

$$\leq \frac{p^2}{4(p-1)} \int(u(y)-u(x))(u^{p-1}(y)-u^{p-1}(x))m_t(dx \times dy)$$

for all $u \in \mathcal{U} \cap D(A)$ and $t > 0$; and certainly this will follow if we show that $(\eta^{p/2}-\xi^{p/2})^2 \leq \frac{p^2}{4(p-1)}(\eta-\xi)\,(\eta^{p-1}-\xi^{p-1})$ for all $0 < \xi < \eta$; or equivalently that $4/p^2\,(1-\eta^{p/2})^2 \leq \frac{1}{p-1}(1-\eta)(1-\eta^{p-1})$ for all $0 < \eta < 1$. But $4/p^2\,(1-\eta^{p/2})^2 = (\int_\eta^1 t^{p/2-1}dt)^2 \leq (1-\eta)\int_\eta^1 t^{p-2}dt = \frac{1}{p-1}(1-\eta)(1-\eta^{p-1})$. □

We can now prove the following slight variation on one of Gross's basic results.

(9.10) <u>Theorem</u>: Assume that $J \underset{m}{\leq} \alpha J_\sigma$. Then for all $p \in (1,\infty)$ and $u \in \mathcal{U} \cap D(A)$:

$$(9.11) \qquad \int u^p \log u\,dm \le \frac{\alpha p}{4(p-1)}(u^{p-1},-Au)_{L^2(m)} + \|u\|^p_{L^p(m)} \log\|u\|_{L^p(m)} .$$

Moreover, if $t > 0$, $p \in (1,\infty)$, and $q_\alpha(t,p) = 1 + (p-1)e^{4t/\alpha}$ then

$$(9.12) \qquad \|P_t\phi\|_{L^q(m)} \le \|\phi\|_{L^p(m)} , \quad \phi \in C_b(E) \text{ and } 1 \le q \le q_\alpha(t,p)$$

Conversely, if, for some $\alpha > 0$ and $p \in (1,\infty)$, $\|P_t\phi\|_{L^{q_\alpha(p,t)}(m)} \le \|\phi\|_{L^p(m)}$, $t > 0$ and $\phi \in C_b(E)$, then $J_m \le \frac{\alpha p^2}{4(p-1)} J_\sigma$. In particular, if $p = 2$, then $J_m \le \alpha J_\sigma$.

<u>Proof</u>: Suppose $J_m \le \alpha J_\sigma$. Then, for all $f \in L^1(m)^+$ with $\|f\|_{L^1(m)} = 1$, $\int f\log f dm \le \alpha\varepsilon(f^{1/2},f^{1/2})$. Thus, for any $f \in L^1(m)^+$,

$$\int f\log f dm \le \alpha\varepsilon(f^{1/2},f^{1/2}) + \|f\|_{L^1(m)} \log\|f\|_{L^1(m)} .$$

In particular, if $u \in \mathfrak{U} \cap D(A)$ and $f = u^p$, then

$$p\int u^p \log u dm \le \alpha\varepsilon(u^{p/2},u^{p/2}) + p\|u\|^p_{L^p(m)} \log\|u\|_{L^p(m)} .$$

Since $\varepsilon(u^{p/2},u^{p/2}) \le \frac{p^2}{4(p-1)}(u^{p-1},-Au)_{L^2(m)}$, this proves (9.11) .

To prove (9.12) , it suffices to show that

$\|P_t\phi\|_{L^{q_\alpha(t,p)}(m)} \le \|\phi\|_{L^p(m)}$, $\phi \in C_b(E)$. Indeed, for any $p \in [1,\infty)$,
$\|P_t\phi\|_{L^1(m)} \le \|\phi\|_{L^1(m)} \le \|\phi\|_{L^p(m)}$; and so
$\|P_t\phi\|_{L^q(m)} \le \|P_t\phi\|^\theta_{L^1(m)} \|P_t\phi\|^{1-\theta}_{L^{q_\alpha(t,p)}(m)} \le \|\phi\|_{L^p(m)}$ if $1/q = \theta + (1-\theta)/q_\alpha(t,p)$,
and $\|P_t\phi\|_{L^{q_\alpha(t,p)}(m)} \le \|\phi\|_{L^p(m)}$. Now, let $u \in \mathfrak{U} \cap D(A)$ be given and set

$u_t = P_t u$. Then, with $q(t) = q_\alpha(t,p)$, we have

$$\frac{d}{dt} \log \|u_t\|_{q(t)}$$

$$= -\frac{q'(t)}{q(t)} \log \|u_t\|_{q(t)}$$

$$+ \frac{q'(t)}{q(t)} \frac{1}{\|u(t)\|_{q(t)}^{q(t)}} \int (u_t)^{q(t)} \log u_t dm$$

$$+ \frac{1}{\|u_t\|_{q(t)}^{q(t)}} \int (u_t)^{q(t)-1} Au_t dm \quad ;$$

and so, (9.11) implies that:

$$\|u_t\|_{q(t)}^{q(t)-1} \frac{d}{dt} \|u_t\|_{q(t)}$$

$$= \frac{q'(t)}{q(t)} \left(\int (u_t)^{q(t)} \log u_t dm - \|u_t\|_{q(t)}^{q(t)} \log \|u_t\|_{q(t)} \right)$$

$$- (u_t^{q(t)-1}, -Au_t)_{L^2(m)}$$

$$\leq \left(\frac{q'(t)\alpha}{4(q(t)-1)} - 1 \right) (u_t^{q(t)-1}, -Au_t)_{L^2(m)} = 0 \quad .$$

In other words, $t \to \|u_t\|_{q(t)}$ is non-increasing. Since $\|u_0\|_{q(0)} = \|u\|_p$, this proves that $\|P_t u\|_{L^{q_\alpha(t,p)}(m)} \leq \|u\|_{L^p(m)}$ for all $t > 0$ and $u \in \mathcal{U} \cap D(A)$ and therefore for all $u \in C_b(E)^+$. Since $|P_t \phi| \leq P_t |\phi|$, $\phi \in C_b(E)$, this completes the proof of (9.12) from $J_m \leq \alpha J_\sigma$.

The converse assertion follows easily from the fact that (9.6) implies (9.1) with $\alpha = \alpha(T,p,q)$. Indeed, use this with $q = q_\alpha(T,p)$ and pass to the limit as $T \downarrow 0$.

$\square$

We now use Theorem (9.10) to give a proof of Nelson's hypercontraction result. Our proof follows the idea of L. Gross.

(9.13) <u>Lemma</u>: Let $\{P^{\ell}_{x_\ell} : x_\ell \in E_\ell\}$, $\ell = 1,2$, satisfy (S.C.) with $m_\ell \in \mathcal{M}_1(E_\ell)$. Denote by $P_\ell(t,x_\ell,\cdot)$ the transition function, by $\{P^{\ell}_t : t > 0\}$ the semigroup on $C_b(E_\ell)$, and by ε_ℓ the Dirichlet form (cf. (7.36)) associated with $\{P^{\ell}_{x_\ell} : x_\ell \in E_\ell\}$. Set $P_x = P^1_{x_1} \times P^2_{x_2}$ for $x = (x_1,x_2)$ $E_1 \times E_2 \equiv E$. Then $\{P_x : x \in E\}$ satisfies (S.C.) with $m = m_1 \times m_2$. Moreover, if $P(t,x,\cdot)$, $\{P_t : t > 0\}$ and ε are, respectively the transition function, the semigroup on $C_b(E)$ and the Dirichlet form associated with $\{P_x : x \in E\}$, then $P(t,x,\cdot) = P(t,x_1,\cdot) \times P(t,x_2,\cdot)$, $P_t = P^1_t \otimes P^2_t$, and

$$\varepsilon(\phi,\phi) \geq \tilde{\varepsilon}_1(\phi,\phi) + \tilde{\varepsilon}_2(\phi,\phi) \quad , \quad \phi \in L^2(m) \quad , \tag{9.14}$$

where $\tilde{\varepsilon}_1(\phi,\phi) \equiv \int_{E_2} \varepsilon_1(\phi(\cdot,x_2),\phi(\cdot,x_2))m_2(dx_2)$ and $\tilde{\varepsilon}_2(\phi,\phi) \equiv \int_{E_1} \varepsilon_2(\phi(x_1,\cdot),\phi(x_1,\cdot))m_1(dx_1)$. Finally, if J^{ℓ}_{σ} is associated with $\{P^{\ell}_{x_\ell} : x_\ell \in E_\ell\}$ and $J_{m_\ell} \leq \alpha J^{\ell}_{\sigma}$ for $\ell = 1,2$, then $J_m \leq \alpha J_\sigma$ where J_σ is associated with $\{P_x : x \in E\}$.

<u>Proof</u>: The facts that $\{P_x : x \in E\}$ satisfies (S.C.) with m and that $P(t,x,\cdot) = P_1(t,x_1,\cdot) \times P_2(t,x_2,\cdot)$ are extremely elementary and their verification is left as an exercise. Clearly the equation $P_t = P^1_t \otimes P^2_t$ follows from the one for the transition functions.

To prove (9.14) , first note that $\overline{P}_t = \overline{P}^1_t \otimes \overline{P}^2_t$ follows immediately from $P_t = P^1_t \otimes P^2_t$. Next, set $\mathcal{F} = \text{span}\{\phi_1 \otimes \phi_2 : \phi_1 \in D(\overline{A}_1)$ and $\phi_2 \in D(\overline{A}_2)\}$. Then $\mathcal{F} \subseteq D(\overline{A}\ )$ and

$\text{graph}(\overline{A}|_{\mathscr{F}})$ is dense in graph $(\overline{A})$. Indeed, if $\phi_1 \in D(\overline{A}_1)$ and $\phi_2 \in D(\overline{A}_2)$, then

$$\frac{1}{t}(\overline{P}_t(\phi_1 \otimes \phi_2) - \phi_1 \otimes \phi_2) = \frac{1}{t}(\overline{P}^1_t\phi_1 \otimes \overline{P}^2_t\phi_2 - \phi_1 \otimes \phi_2)$$

$$\rightarrow \overline{A}_1\phi_1 \otimes \phi_2 + \phi_1 \otimes \overline{A}_2\phi_2 \quad .$$

Thus $\phi_1 \otimes \phi_2 \in D(\overline{A})$ and

(9.15) $$\overline{A}(\phi_1 \otimes \phi_2) = \overline{A}_1\phi_1 \otimes \phi_2 + \phi_1 \otimes \overline{A}_2\phi_2 \quad .$$

In particular, $\mathscr{F} \subseteq D(\overline{A})$. Also, given $\phi \in D(\overline{A})$ and $\varepsilon > 0$, choose $0 < t \leq 1$ so that $\|\phi-\overline{P}_t\phi\|_{L^2(M)} + \|\overline{A}\phi-\overline{AP}_t\phi\|_{L^2(m)} < \varepsilon$. Next, choose $\psi \in \text{span}\{\psi_1 \otimes \psi_2 : \psi_1 \in L^2(m_1) \text{ and } \psi_2 \in L^2(m_2)\}$ so that $\|\phi-\psi\|_{L^2(m)} < t\varepsilon$. Since $\|\overline{AP}_t\|_{\text{Hom}(L^2(m);L^2(m))} \leq 1/t$, $\|\overline{P}_t\phi-\overline{P}_t\psi\|_{L^2(m)} + \|\overline{AP}_t\phi-\overline{AP}_t\psi\|_{L^2(m)} < 2\varepsilon$. Since $\overline{P}_t\psi \in \mathscr{F}$, this completes the proof that graph $(\overline{A}|_{\mathscr{F}})$ is dense in graph $(\overline{A})$.

From (9.15) , it is clear that

$$\varepsilon(\phi,\phi) = \tilde{\varepsilon}_1(\phi,\phi) + \tilde{\varepsilon}_2(\phi,\phi)$$

for $\phi \in \mathscr{F}$. But, since $\text{graph}(\overline{A}|_{\mathscr{F}})$ is dense in graph $(\overline{A})$, if $\phi \in L^2(m)$ satisfies $\varepsilon(\phi,\phi) < \infty$, then there exists $\{\phi_n\}_1^\infty \subseteq \mathscr{F}$ such that $\|\phi-\phi_n\|_{L^2(m)} + \varepsilon(\phi-\phi_n,\phi-\phi_n)^{1/2} \rightarrow 0$ as $n \rightarrow \infty$. Since, by Fatou's lemma, $\tilde{\varepsilon}_\ell(\phi,\phi) \leq \overline{\lim_{n\to\infty}}\, \tilde{\varepsilon}_\ell(\phi_n,\phi_n)$, $\ell = 1,2$, (9.14) now follows.

Finally, assume that $J_{m_\ell} \leq \alpha J^\ell$, $\ell = 1,2$. Then, for all $\phi_\ell \in L^2(m_\ell)$:

$$\int \phi_\ell^2 \log \phi_\ell \, dm_\ell \le \frac{\alpha}{2}\, \varepsilon_\ell(\phi_\ell,\phi_\ell) + \|\phi_\ell\|^2_{L^2(m_\ell)} \log \|\phi_\ell\|_{L^2(m_\ell)} \quad .$$

In particular, if $u \in \mathcal{U} \cap D(A)$, then, for each $x_2 \in E_2$:

$$\int u(x_1,x_2)^2 \log u(x_1,x_2) m_1(dx_1)$$

$$\le \frac{\alpha}{2}\, \varepsilon_1(u(\cdot,x_2),u(\cdot,x_2))$$

$$+ \|u(\cdot,x_2)\|^2_{L^2(m_1)} \log \|u(\cdot,x_2)\|_{L^2(m_1)} \quad ;$$

and so:

$$\int u^2 \log u \, dm \le \frac{\alpha}{2}\, \tilde{\varepsilon}_1(u,u)$$

$$+ \frac{\alpha}{2}\, \varepsilon_2((\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2},(\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2})$$

$$+ \|u\|^2_{L^2(m)} \log \|u\|_{L^2(m)} \quad .$$

We must next show that

(9.16)
$$\varepsilon_2((\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2},(\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2}) \le \tilde{\varepsilon}_2(u,u) \quad .$$

To this end, define $(\varepsilon_2)_t$ and $(m_2)_t$ for $t > 0$ as in Lemma (7.38) . Then

$$\varepsilon_2((\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2},(\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2})$$

$$= \lim_{t \downarrow 0} (\varepsilon_2)_t((\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2},(\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2}) \quad ;$$

while, for each $x_1 \in E_1$ and $t > 0$:

$$(\varepsilon_2)_t(u(x_1,\cdot),\ u(x_1,\cdot)) \le \varepsilon_2(u(x_1,\cdot),u(x_1,\cdot)) \quad .$$

Hence, (9.16) will be proved once we check that

$$(\varepsilon_2)_t((\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2},(\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2})$$
$$\le \int_{E_1} (\varepsilon_2)_t(u(x_1,\cdot),u(x_1,\cdot))m_1(dx_1)$$

for each $t > 0$. But:

$$t(\varepsilon_2)_t((\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2},(\int u(x_1,\cdot)^2 m_1(dx_1))^{1/2})$$

$$= \int\int_{E_2\times E_2} (\|u(\cdot,y_2)\|_{L^2(m_1)} - \|u(\cdot,x_2)\|_{L^2(m_1)})^2 (m_2)_t(dx_2 \times dy_2)$$

$$\le \int\int_{E_2\times E_2} \|u(\cdot,y_2) - u(\cdot,x_2)\|^2_{L^2(m_1)} (m_2)_t(dx_2 \times dy_2)$$

$$= \int_{E_1} (\int\int_{E_2\times E_2} (u(x_1,y_2)-u(x_1,x_2))^2 (m_2)_t(dx_2 \times dy_2))m_1(dx_1)$$

$$= \int_{E_1} (\varepsilon_2)_t(u(x_1,\cdot),u(x_1,\cdot))m_1(dx_1)$$

Thus, (9.16) has been proved. Using (9.16) and (9.14) in the equation which precedes it, we now see that:

$$\int u^2 \log u\, dm \le \frac{\alpha}{2}\, \varepsilon(u,u) + \|u\|^2_{L^2(m)} \log\|u\|_{L^2(m)}$$

for all $u \in \mathcal{U} \cap D(A)$. Finally, given $f \in L^1(m)^+$ satisfying $\|f\|_{L^1(m)} = 1$ and $\varepsilon(f^{1/2},f^{1/2})$, we can use Lemma (8.20) to find $\{u_n\}_1^\infty \subseteq \mathcal{U} \cap D(A)$ such that $\|u_n\|_{L^2(m)} = 1$, $u_n \to f^{1/2}$ in $L^2(m)$, and $\varepsilon(u_n,u_n) \to \varepsilon(f^{1/2},f^{1/2})$. Since $\xi^2\log\xi = 1/2\ \xi^2\log\xi^2 \ge e^{-1}/2$ for $\xi \ge 0$, Fatou's lemma allows us to

conclude that:

$$1/2\int f\log f dm \leq \varliminf_{n\to\infty} \int u_n^2 \log u_n \leq \frac{\alpha}{2} \varliminf_{n\to\infty} \varepsilon(u_n,u_n)$$

$$= \frac{\alpha}{2} \varepsilon(f^{1/2},f^{1/2}) \quad .$$

Clearly, $J_m \leq \alpha J_\sigma$ follows immediately from this.

□

(9.17) Exercise: Let $\mu(dt) = e^{-t}dt$ on $(0,\infty)$ and set $Q = \mu^{Z^+}$ on R^{Z^+}. Define $N(t)$, for $t \geq 0$, by $N(t) = \max\{n \geq 0 : \sum_{i=0}^{n} \tau_i \leq t\}$ ($\sum_{i=1}^{0} \tau_i \equiv 0$), where $\tau_i : R^{Z^+} \to (0,\infty)$ is the $i^{\underline{th}}$ coordinate map. Note that $t \to N(t)$ is a non-decreasing, right continuous map and show that

$$Q(N(t)-N(s) = n \,|\, \sigma(N(u) : 0 \leq u \leq s)) = \frac{(t-s)^n}{n!} e^{-(t-s)}$$

for all $0 \leq s < t$ and $n \geq 0$. Next, define $Y(t,x)$, $t \geq 0$ and $x \in \{-1,1\}$, by $Y(t,x) = x(-1)^{N(t/2)}$ and denote by Q_x the measure $Q \circ Y(\cdot,x)^{-1}$ on $D([0,\infty);\{-1,1\})$. Show that $\{Q_x : x \in \{-1,1\}\}$ is a Feller continuous, time homogeneous Markov family with transition function $Q(t,x,\cdot)$ given by:

$$1 - Q(t,x,\{-x\}) = Q(t,x,\{x\}) = \frac{1+e^{-t}}{2} \quad .$$

In particular, if $\phi(x) = a+bx$, $x \in \{-1,1\}$, check that the associated semigroup $\{Q_t : t > 0\}$ acts on ϕ by $[Q_t\phi](x) = a + e^{-t}bx$. Next, let $m \in \mathcal{M}_1(\{-1,1\})$ be given by $m(\{1\}) = m(\{-1\}) = 1/2$, check that $\{Q_x : x \in \{-1,1\}\}$ satisfies (S.C.) with respect to m. Finally, show that if ϕ is given as above, then the associated Dirichlet form ε acts on ϕ by $\varepsilon(\phi,\phi) = b^2$.

(9.18) <u>Lemma</u>: Let $\{Q_x : x \in \{-1,1\}\}$, $\{Q_t : t > 0\}$ and m be as in (9.17) . For each $n \geq 1$, let $Q_t^n = (Q_t)^{\otimes n}$ on $C((\{-1,1\})^n)$ and $m_n = (m)^n$ on $(\{-1,1\})^n$. Then, for each $p \in (1,\infty)$, all $t > 0$, and $1 \leq q \leq 1 + (p-1)e^{2t}$, $\|Q_t^n\phi\|_{L^q(m_n)} \leq \|\phi\|_{L^p(m_n)}$, $\phi \in C((\{-1,1\})^n)$.

<u>Proof</u>: In view of Lemma (9.13) and Theorem (9.10) , all that we have to show is that $J_m \leq 2J_\sigma$, where J_σ is associated with $\{Q_x : x \in E\}$. Clearly, this is equivalent to showing that if $\phi \in L^2(m)^+$, then $\int\phi^2\log\phi dm \leq \varepsilon(\phi,\phi) + \|\phi\|^2_{L^2(m)} \log\|\phi\|_{L^2(m)}$. Note that the most general $\phi : \{-1,1\} \to [0,\infty)$ has the form $\phi(x) = a + bx$, where $|b| \leq a$. Since $\varepsilon(\phi,\phi) = b^2$, there is nothing to prove when $a = 0$. Thus, by homogeneity, we need only look at $\phi_b(x) = 1 + bx$, where $|b| \leq 1$. In fact, by symmetry, we can and will restrict ourselves to $0 \leq b \leq 1$. That is, we need only show that:

$$\begin{aligned}
h(b) &\equiv (1+b)^2\log(1+b) + (1-b)^2\log(1-b) - (1+b^2)\log(1+b^2) \\
&= 2[\int\phi_b^2\log\phi_b dm - \|\phi_b\|^2_{L^2(m)} \log\|\phi_b\|_{L^2(m)}] \\
&\leq 2\ \varepsilon(\phi_b,\phi_b) = 2b^2
\end{aligned}$$

for $0 \leq b \leq 1$. Note that $h(0) = 0$. Also,

$$\begin{aligned}
h'(b) &= 2[(1+b)\log(1+b) + 1/2(1+b) - (1-b)\log(1-b) - 1/2(1-b) \\
&\quad - b\log(1+b^2) - b] \\
&= 2[(1+b)\log(1+b) - (1-b)\log(1-b) - b\log(1+b^2)] \ .
\end{aligned}$$

In particular, $h'(0) = 0$. Finally,

$$h''(b) = 2[\log(1+b) + 1 + \log(1-b) + 1 - \log(1+b^2) - \frac{2b^2}{1+b^2}]$$

$$= 4 + 2\log((1-b^2)/(1+b^2)) - \frac{4b^2}{1+b^2} \le 4 \quad .$$

Thus:

$$h(b) = \int_0^b d\tau \int_0^\tau h''(\sigma)d\sigma \le 4\,\frac{b^2}{2} = 2b^2 \quad .$$

$\square$

(9.19) <u>Theorem</u> (L. Gross and E. Nelson): Let $\{P_x : x \in R^1\}$ be the Ornstein-Uhlenbeck family described in (8.28), and denote by $\{P_t : t > 0\}$ the associated semigroup. Let $\gamma(dy) = \frac{1}{(2\pi)^{1/2}} e^{-y^2/2}dy$. Then $J_\gamma \le J_\sigma$; and, for $p \in (1,\infty)$, $t > 0$, and $1 \le q \le 1 + (p-1)e^t$,

$$\|P_t\phi\|_{L^q(\gamma)} \le \|\phi\|_{L^p(\gamma)} \quad , \quad \phi \in C_b(E) \quad .$$

<u>Proof</u>: Define $\{Q_x : x \in \{-1,1\}\}$ and m as in (9.17) and let μ_t on R^2 denote the distribution of $\binom{x(0)}{x(t)}$ under $Q_m = \int Q_x m(dx)$. Then, for $\phi,\psi \in C_b(R^1)$: $\int_{R^1}\int_{R^1} \phi(\xi)\psi(\eta)\mu_t(d\xi\times d\eta) = \int_{R^1}\phi(\xi)[Q_t\psi](\xi)m(d\xi)$. In particular, μ_t has mean $\binom{0}{0}$ and covariance $A(t) \equiv \begin{pmatrix} 1 & e^{-t} \\ e^{-t} & 1 \end{pmatrix}$. Thus, by the central limit theorem, if $\Phi_n : (R^2)^n \to R^2$ is defined by $\Phi(\binom{\xi_1}{\eta_1},\ldots,\binom{\xi_n}{\eta_n}) = \frac{1}{n^{1/2}}\sum_1^n \binom{\xi_i}{\eta_i}$, then $\mu_t^n \circ \Phi_n^{-1} \Rightarrow \Gamma_{A(t)}$, where $\Gamma_{A(t)}$ is the Gaussian measure on R^2 with mean $\binom{0}{0}$ and covariance $A(t)$. In particular, if $\phi,\psi \in C_b(R^1)$, then

$$\int_{R^1}\int_{R^1}\phi(x)\psi(y)\Gamma_{A(t)}(dx \times dy)$$

$$= \lim_{n\to\infty}\int_{R^n}\int_{R^n}\phi\left(\frac{\xi_1+\cdots+\xi_n}{n^{1/2}}\right)\psi\left(\frac{\eta_1+\cdots+\eta_n}{n^{1/2}}\right)\mu_t^n(d\xi \times d\eta) \quad .$$

Next, note that for $\Phi,\Psi \in C_b(R^n)$:

$$\int_{R^n}\int_{R^n}\Phi(\xi)\Psi(\eta)\mu_t^n(d\xi \times d\eta) = \int_{R^n}\Phi(\xi)Q_t^n\Psi(\eta)m_n(d\xi \times d\eta) \quad .$$

Thus, by Lemma (9.18) :

$$\left|\int_{R^n}\int_{R^n}\phi\left(\frac{\xi_1+\cdots+\xi_n}{n^{1/2}}\right)\psi\left(\frac{\eta_1+\cdots+\eta_n}{n^{1/2}}\right)\mu_t^n(d\xi \times d\eta)\right|$$

$$\leq \left(\int_{R^n}\left|\phi\left(\frac{\xi_1+\cdots+\xi_n}{n^{1/2}}\right)\right|^{q'} m_n(d\xi)\right)^{1/q'}\left(\int_{R^n}\left|\psi\left(\frac{\eta_1+\cdots+\eta_n}{n^{1/2}}\right)\right|^{p} m_n(d\eta)\right)^{1/p}$$

when $\frac{1}{q'} = 1 - \frac{1}{q}$ and $1 \leq q \leq 1 + (p-1)e^{2t}$. Applying the central limit theorem once again, we now see that:

$$\left|\int_{R^1}\int_{R^1}\phi(x)\psi(y)\Gamma_{A(t)}(dx \times dy)\right| \leq \|\phi\|_{L^{q'}(\gamma)}\|\psi\|_{L^p(\gamma)} \quad .$$

Finally,

$$\Gamma_{A(t)}(dx \times dy) = \frac{1}{2\pi(1-e^{-2t})^{1/2}}\exp\left(-\frac{(y^2e^{-2t}-2xye^{-t}+x^2e^{-2t})}{2(1-e^{-2t})}\right)dxdy$$

$$= P(2t,x,dy)\gamma(dx)$$

where $P(t,x,\cdot)$ is the transition function for the Ornstein-Uhlenbeck family. Thus the preceding becomes

$$\left|\int_{R^1}\phi\cdot P_{2t}\psi d\gamma\right| \leq \|\phi\|_{L^{q'}(\gamma)}\|\psi\|_{L^p(\gamma)}$$

for all $\phi, \psi \in C_b(E)$. Clearly this implies that $\|P_{2t}\phi\|_{L^q(\gamma)} \leq \|\phi\|_{L^p(\gamma)}$, $\phi \in C_b(E)$, $t > 0$, $p \in (1,\infty)$, and $q \leq 1 + (p-1)e^{2t}$. The relation $J_m \leq J_\sigma$ now follows from the last part of Theorem (9.10) with $p = 2$. $\square$

(9.20) <u>Exercise</u>: Referring to the situation in Theorem (9.19) , show that for each $p \in (1,\infty)$, $t > 0$, and $q > 1 + (p-1)e^t$: $\sup\{\|P_t\phi\|_{L^q(\gamma)} : \phi \in C_b(E)$ and $\|\phi\|_{L^p(\gamma)} = 1\} = \infty$. In other words, Theorem (9.19) is sharp.

[D. & V., Oxford] Donsker, M. and Varadhan, S.R.S., "Asymptotic evaluation of certain Wiener integrals for large time," in Functional Integration and its Applications (ed. by A.M. Arthur), Oxford Press, pp. 15-33 (1975).

[D. & V., Parts I & III] Donsker, M. and Varadhan, S.R.S., "Asymptotic evaluation of certain Markov process expectations for large time," I and III, Comm. Pure Appl. Math. #28 and 29, pp. 1-47 and pp. 389-461, (1975) and (1976)

[Kac] Kac, M., Integration in Function Spaces and Some of Its Applications, 1980 Fermi Lectures, Publ. by Academia Nazionale dei Lincei, Scuola Normale Superiore.

[Fnal. Anal., L., Schwartz] Functional Analysis, C.I.M.S. - N.Y.U. Lecture Notes.

[S. & V.] Stroock, D. and Varadhan, S.R.S. Multidimensional Diffusion Processes, Springer-Verlag Grund. Series #233, 1979.

[S., St. Flour] Stroock, D., "Some applications of stochastic calculus to partial differential equations," Springer-Verlag Lec. Notes in Math #976, ed by P.L. Hennequin, pp. 267-382, 1983.